On the Regulation of Central Carbon Metabolism in *S. cerevisiae*

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)
im Fach Biophysik

eingereicht an der
Mathematisch-Naturwissenschaftlichen
Fakultät I
Humboldt-Universität zu Berlin

von
Herrn mag. rer. nat. József Bruck

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen
Fakultät I:
Prof. Dr. Andreas Herrmann

Gutachter:

1. Prof. Dr. Dr. h.c. Edda Klipp
2. Prof. Dr. Hermann-Georg Holzhütter
3. Juha-Pekka Pitkänen, Ph.D.

eingereicht am:	26. Okt. 2011
Tag der mündlichen Prüfung:	4. Mai 2012

Abstract

In this work, we aimed to elucidate central carbon metabolism focusing on the aspect of regulation, especially by separating two regulatory levels: metabolic regulation, associated with direct interactions of metabolites and enzymes, and hierarchic regulation, associated with enzyme level change via regulation of *de novo* enzyme production. Our investigations were largely based on the analysis of three datasets from glucose limited continuous cultures of *S. cerevisiae* with five different oxygen provision levels ranging from anaerobic to highly aerobic. These datasets contained data on transcript levels, intracellular metabolite levels, and intracellular flux distributions, respectively. The experiments were performed by collaborators.

Extracellular conditions on the macroscopic scale were investigated in Chapter 2. This was inspired by the perceived lack of clarity regarding an important aspect: concentration of glucose, the limiting nutrient and main carbon source in these cultures. The main outcome of this theoretical analysis was characterisation of the selection pressure in a chemostat culture, as selecting for cells (or metabolic states) which produce the growth rate, defined by the pre-set dilution rate, with lower external concentration of the limiting nutrient.

Flux regulation on the scale of individual enzymes was investigated for selected reactions, and the phosphofructokinase - fructobisphosphatase unit in Chapter 3. This analysis was based on the attempt to reproduce flux changes through these reactions, using enzyme kinetic expressions with inputs from the three aforementioned datasets. The notion of hierarchic and metabolic regulation was introduced and modified.

System-level analysis of central carbon metabolism was undertaken in Chapter 4. Using the information on metabolite levels and flux, a kinetic model representing significant parts of central carbon metabolism was constructed. Most kinetic expressions in the model were taken from a pre-existing kinetic model (Teusink model). In order to arrive at feasible flux distributions, constrained metabolic flux balance analysis was performed, using a stoichiometric network, constructed to be consistent with the model's stoichiometry.

Fitting the model resulted in two sets of parameters corresponding to steady states reproducing, the nominal data values of the anaerobic and the fully aerobic conditions.

Finally, an *in silico* perturbation experiment, mimicking the sudden introduction of oxygen into the system, was performed. No further modification or fitting of the model was undertaken for this purpose.

Zusammenfassung

Ziel dieser Arbeit war es, den zentralen Kohlenstoffwechsel mit besonderem Fokus auf Regulation zu untersuchen, insbesondere durch die Auftrennung von zwei Regulationsebenen: metabolische Regulation, assoziiert mit direkten Wechselwirkungen zwischen Metaboliten und Enzymen, sowie hierarchische Regulation, assoziiert mit Änderungen in Enzymmengenänderungen durch die Regulation von *de novo* Enzymproduktion. Unsere Untersuchungen basieren größtenteils auf drei Datensätzen aus glukoselimitierten Chemostatkulturen von *S. cerevisiae* mit fünf verschiedenen Sauerstoffversorgungsstufen von anaerob bis hochaerob. Diese Datensätze enthielten Daten über Transkription, intrazelluläre Metabolitengehalt und intrazelluläre Flußverteilung. Die entsprechenden Experimente wurden durch Kollegen ausgeführt.

Im *Kap. 2* wurden Extrazelluläre Bedingungen im Makroskopischen untersucht. Dies wurde durch die Unklarheit in Bezug auf einen wichtigen Aspekt, die Konzentration von Glukose, der limitierende Nährstoff und Hauptkohlenstoffquelle in diesen Kulturen inspiriert. Das wichtigsten Ergebnis dieser theoretischen Analyse ist die Charakterisierung des Selektionsdruckes in einem Chemostatkultur, nämlich dadurch, dass die (durch die Verdünnungsrate D festgelegte) Wachstumsrate mit geringster externen Konzentration des limitierende Nährstoffes vonstatten geht.

Im *Kap. 3* wurde Fluss und Regulation der einzelnen Enzyme für ausgewählte Reaktionen sowie für die Phosphofructokinase - Fructobisphosphatase Regelungseinheit untersucht. Diese Analyse wurde auf den Versuch basiert, Flussänderungen durch diese Reaktionen mithilfe von enzymkinetischen Ausdrücke sowie mit Inputs aus den drei genannten Datensätzen, zu reproduzieren. Der Begriff der hierarchischen und der metabolischen Regulation wurde eingeführt und angepaßt.

Im Kap. 4 wurde eine Analyse auf Systemebene des zentralen Kohlenstoffwechsels durchgeführt. Unter Verwendung der Metaboliten- und der Flußdaten wurde ein kinetisches Modell konstruiert, welches wesentliche Teile des zentralen Kohlenstoffwechsels umfaßt. Die meisten kinetischen Ausdrücke und Parameterwerte wurden aus einem bestehenden kinetischen Modells (Teusink-Modell) übernommen. Um zulässige Flussverteilungen zu erhalten, welches als Gleichgewichtsfluß des Modells dienen kann, mußten wir eine Flußbalanceanalyse mit Zwangsbedingungen durchführen.

Fitting des Modells an die Daten führte zu Parametersätzen welche mit dem anaeroben, bzw. aeroben stationären Zustand assoziiert werden. Abschließend wurde eine *in silico* Perturbationsexperiment durchgeführt, welches das plötzliche einführen von Sauerstoff in das System nachahmt.

Dedication

Drága Nagypapának,
e para Ciça, meu amor,
and to the rest of my family in between!

We offer up our predictions to our Lord, Consistency,
Who is holding The Instrument in his left
Who is wielding Occam's Razor in his right
Who goes to rest in the shade of *Unentscheidbarkeit*

- from the theoretician's morning choral -

Contents

Chapter 1

An introduction

*L*IFE COPES - an observation, so readily made within the biosphere of our planet that it was seldom thought to be surprising to see the same organism cope with a variety of conditions, and to find life in a rather wide range of conditions. Coping with the changes of environment seems a fundamental property of life.

In a post-Darwinian world, the above may be rephrased as 'Life is what has coped (prevailing over others in the past)'. While evolutionary theories offer an appealing conceptual framework for the history of life, as well as view thereof as emergent behaviour of certain systems composed of non-alive components, it is not clear at this point, to what extent they are able to explain, let alone, predict details. Attempts have been made to give a reasonably detailed account of bioevolution [81, Smith and Szathmáry].

While no consensus on a definition of 'life' is available, metabolism - the set of processes involved with material flow through an organism - is widely regarded as a key property of living systems. This opinion is also mirrored in a theoretical model of life, the 'chemoton' [30] (see description in [81]). This implies that viruses, entities lacking metabolism, and in a seemingly direct implication incapable of replication using only 'their own' body, are not regarded as 'alive'. On the other hand, they are clearly an active part of their ecosystem, shaping its fate through their evolution.

It follows from the above that metabolism provides all organisms with material to maintain, and to replicate their bodies. Heterotroph and lithotroph organisms also gain the necessary free energy from the inflow of material, while autothrophs use light. Since litotrophs are a relative minority today, the main inflow of free energy into our biosphere is provided by sunlight via autotroph organisms.

As a classical attribute of living systems, main parts of metabolism belonged to the first fields within molecular biology to be elucidated. Much of biochemistry evolved by elucidating those properties of glycolysis and the TCA-cycle which are 'textbook knowledge' today.

However, the statement in the beginning is not necessarily connected to metabolism. Notably, it seems to be so self-evident that the above mentioned authors seem to have found it unworthy of mentioning: response to external stimuli in a way directed by self preservation. While life without metabolism is difficult to imagine, life without the response to external stimuli seems impossible.

Hence, metabolism exhibits regulation: the flow of material[1] is modified in response to changes in environmental conditions.

Interestingly, even though metabolism is counted among the best understood parts of cellular biology by many, regulation thereof is far from being well understood. We seem to just have begun putting together the first pieces of the puzzle. Perhaps the most complete such piece is metabolism at the level of enzymes and metabolites. Enzyme kinetics provides, in principle, a conceptual connection between physics and biology.

This work attempts to contribute to the understanding of the regulation of central carbon metabolism, the main distributor of material and free energy in the cell - here, the latter is the case, too, since the observations we refer to were made on the species *Saccharomyces cerevisiae*, a heterotrophic organism. While this thesis reports on purely theoretical work, we attempted to gain as good understanding of the experimental system as possible. The presentation is divided into three major chapters.

The *second chapter* introduces and analyses the experimental system known as glucose limited chemostat cultivation, and offers short descriptions of the measurement techniques, as well as of some characteristics of the resulting datasets. The three datasets contain information about intermediate metabolite levels, intermediate metabolite fluxes, and transcription activities.

The *third chapter* presents a step in the integrative analysis of the introduced datasets in combination with enzyme kinetics equations. The latter may be interpreted as the integration of existing knowledge about the involved enzymes - e.g. regarding parameter values - but also the means to test the inclusion of allosteric interactions.

The main question posed in this chapter is, to which extent the available information is consistent with the measured flux changes through single metabolic reactions. This approach is based on the view that chemical reaction steps catalysed by enzymes are natural building blocks of metabolic pathways.

In the *fourth chapter*, we change the focus of our investigation to a larger scale: we will consider pathways making up a significant portion of central carbon metabolism. We will aim to investigate, to what extent a kinetic model - containing considerable amounts of biochemical knowledge - is consistent with the datasets presented in Chapter 2.

The model which has been constructed for this investigation is a modification, and extension of a published glycolysis model. The model was modified such that, in addition to anaerobic, also respiratory energy metabolism is represented, albeit in a highly simplified way. This makes it possible to represent metabolism in anaerobic as well as aerobic yeast cultures.

In order to provide flux distributions which are consistent with the model's stoichiometry, the constrained flux balance analysis procedure used in producing the internal flux dataset was partly repeated, using a modified stoichiometric network.

The result may be characterised as a kinetic model which is stoichiometrically consistent with a larger stoichiometric network. Beyond the immediate aim of this study, it is presented in the hope to be useful for further studies as well as a

[1] The decision whether photons are regarded as 'material' is left to the reader. The remainder of this text will be concerned with heterotroph organisms.

basis for further extensions.

Finally, the *fifth chapter* provides a summary of the main results, and an extended discussion of some of the present, as well as speculations about future developments.

How far can we dismantle regulation into its components, and what may be the next stations of our understanding? The author hopes that this work contributes, if not to the answers, then to the even more important questions.

Chapter 2

Cell culture and data: basics for an integrative analysis

2.1 Introduction

This chapter is devoted to discussion of the source of data used in this work: the understanding of chemostats cultivations, and the experimental methods used to obtain the data, offering miniature introductions to these techniques. Data from all three datasets are included either here or in Appendix A.

The presented theoretical discussion of the chemostat experimental system may appear more detailed than strictly necessary for the immediate purpose of this work. However, this analysis was inspired by a perceived controversy regarding an important aspect of the experiments: the concentration of glucose, the main nutrient in the cultures determining external conditions for the cells under study. This is intrinsically linked to understanding physiology, a major objective of this work. As a result of this analysis, the existence of selection pressure in such cultures is deduced and characterised. Implications regarding population homogeneity and data reproducibility - in comparison to batch cultures - are discussed.

The second part of this chapter offers a short description of the measurement techniques, and some characteristics of the resulting three datasets. Since these measurements were performed in relatively distinct projects, an integrative discussion is missing. While this work does not aim to serve as a comprehensive integrative discussion of the experimental results, it highlights some aspects, such as the connection between single samples, cultures, and datasets.

2.2 The chemostat: theory of an experimental system

2.2.1 Basic principles and description

As a mean to create growing microbial cultures without temporal changes (inherent to batch cultivations), the experimental concept of continuous cultures was simultaneously developed by the physicist Leó Szilárd who worked with Aaron Novick[1] [62], and by Jacques Monod [58]. The names used by these inventors ('chemostat' and 'continuous culture', respectively) are often used as synonyms. Chemostats fall into the category of stirred bioreactors with continuous operation. For a modern description see, for example, the textbooks [85], [59], or the review [18].

Since experiments in the systems biology community are typically based on more affordable batch cultivations - such as shake flasks - we will highlight differences between these and chemostat cultivations. The motivation behind this is the question, to which extent may datasets based on different cultivations be compared. An answer to this would be valuable in a field in which integration of multiple datasets is practice.. While an absolute answer is not within the scope of this work, it is hoped that the reader will gain insight to help with decisions in particular cases. Why chemostat cultivations can be expected to result in higher reproducibility of data, will be discussed at the end of this section.

At the macroscopic level, the difference between batch and chemostat cultures is characterised by the ability of the latter to provide a *continuous steady state* culture. This implies easier monitoring of many parameters of the culture, such as concentrations, external to the cells. Moreover, one physiological parameter, the specific steady state growth rate of the culture is set independently by the experimenter. Importantly, the chemostat is capable of attaining a steady state (in terms of physiological parameters of the whole culture), a fact corresponding to the stability of the steady state of the equations discussed here.

A core element of the conceptual framework behind our argumentation is the assumptions that a cell in a cultivation only receives information from its microenvironment. Characterising and, if possible, independently setting this mircoenvoronment is one of the main goals of system biology experiment designs. We will furthermore assume, that changes in the cell's microenvironment are fully characterised by changes in concentrations - reflecting our focus on processes involving chemical species, rather than quantities such as temperature of radioactive radiation. Consequently, changes in the 'rest of the world' - such as increasing the glucose influx into the culture - can only influence the cell's physiology by changing concentrations in its microenvironment by mixing or diffusion. We may refer to this as the assumption of no-*fernwirkung*[2] in cell

[1] Both of them had worked on the Manhattan Project - the Hungarian-German-American physicist-turned-biologist Szilárd having conceived the idea of nuclear chain reaction and contributed to the Einstein letter to Roosevelt - but were, like other involved scientists, disgusted by the way the technology was used to end the War, and campaigned against the use of nuclear weapons afterwards.

[2] Even though Einstein's opinion about the 'spooky action-at-a-distance' in the context of the so-called EPR debate [23] seems to have turned out to be false.

cultures.

While, in our view, the above assumptions describe well the majority of experimental scenarios, they are not consistent with commonly encountered arguments based on fluxes into the culture *directly* influencing the cell's physiology, nor with cell growth in a culture with zero concentration of the limiting nutrient.

THE KEY ASSUMPTION, necessary for the outlined construction of a continuous steady state culture is that *the growth rate of a microorganism is a strictly monotonous function of the provided nutrition concentration*[3]. Under circumstances when this condition is not met - e.g. if the growth rate of the culture has reached its maximum - a chemostat may not attain a stable steady state.

A chemostat may be defined as a bioreactor with certain properties which correspond to well-defined experimental concepts. These involve

- (*i*) main volume, required to be *well mixed*
- (*ii*) liquid feed
- (*iii*) limiting nutrient contained in the feed solution
- (*iv*) effluent, specifying the *dilution rate*

The following discussion attempts to elucidate these concepts, following an approach somewhat different from other texts encountered by the author ([18], [96], [85], [59]). Striving to increase conceptual clarity, we will explicitly state some necessary assumptions usually made implicitly.

Design principles

The first object from the above list, the *main (working) volume*, is defined for the purpose of this text as the the liquid phase in the bioreactor, assumed to fully contain the cell culture. Concentrations in the main volume are often called *residual concentrations* (since often measured in the effluent, as discussed below). The main volume is supplied by a *(i) feed*, defined as an incoming flux [mol time^{-1}] of the chosen nutrient composition which is required to be of consant composition resulting in time-independent concentrations and generated influxes of the provided chemical species into the bioreactor.

In addition to the fluid-phase feed, there may be a gaseous feed (controlling, for example, oxygen inflow) usually requiring an exhaust system, to carry away gaseous metabolic products (such as carbon-dioxide).

Since the transformations of the provided chemical species into biomass and products in the bioreactor usually involves mass transfer from or into the gaseous phase, the main volume might vary slightly if the metabolic state of the cells changes. Depending on the particular experimental conditions and the required precision, this can often be neglected.

THE MAIN VOLUME IS REQUIRED to be *(ii) well mixed*, for local measurement values of intensive quantities - such as concentrations, pH, or temperature -

[3] In other words, the growth rate increases if nutrition concentration rises, other experimental conditions assumed being constant.

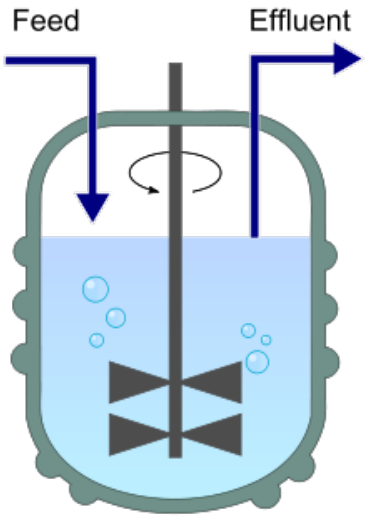

Figure 2.1: Simplified scheme of a chemostat. Of the key concepts, indicated are the mixing, the feed and the effluent. Not indicated but also necessary is the composition of the feed such that a single limiting nutrient is present. Although not required *per se*, gaseous feed and exhaust are necessary in many scenarios to maintain viable conditions.

to describe the whole volume, i.e. the actual microenvironment of each cell. While this property is often associated with complete spatial homogeneity with regard to these quantities, this is hard to fully achieve in reality. This follows from the fact that cells under nutrient limitation are expected to be sensitive to nutrient concentration differences, while at the same time this concentration will be slightly higher near the feed entering the main volume.

However, it turns out that for many purposes it suffices if each cell experiences the same microenvironment *on average*. To formulate this requirement in an accurate way, let's assume that one can follow the trajectory of any small volume of liquid (on the scale of a cell's microenvironment) within the bioreactor, continuously monitoring all relevant intensive quantities in it.

We will regard a bioreactor to be *well mixed* if the following statement holds: Even if these quantities, monitored in the volume, are not fully constant during its trajectory, their time average should attain the macroscopically measured average value faster than the time scale of those processes which we intend to study under homogeneous conditions[4].

For example, if the processes to be studied under homogeneous conditions are assumed to have a characteristic timescale of seconds - as it is the case for many signalling events - it suffices to ensure that each cell experiences the same conditions when averaged on a timescale of 100 miliseconds. In practice, this means that the mixing has to prevent the formation of "still areas" which would mix too slow with the rest of the working volume.

However, the mixing also has another purpose. For meaningful measurements to be made, the mixing must be effective enough that a small sample of volume, taken at a pre-defined location in the reactor, reliably gives the average values

[4] In a more technical - but shorter - formulation, the statistics of each intensive quantity on the ensemble of small fluid volumes is required to be *ergodic* on a time scale below that of the processes of interest.

of the measured quantities in the main volume.

Thus, if the above requirements are fulfilled, we can assume the main volume to appear homogeneous for the purposes of the experiment.

THE FEED IS REQUIRED to supply a *(iii) nutrient solution containing one limiting nutrient*, S_{lim} (e.g. glucose). The underlying notion is that the concentration of S_{lim} in a cell's microenvironment should be the key controlling factor of its physiology while the concentration of other nutrients and products should ideally have negligible effect, at least within the range of the experiment. Another usual formulation of this notion is that all non-limiting nutrients are to be supplied "in excess" compared to the cell's needs per unit of S_{lim} consumed. Hence, the property of a certain nutrient being limiting should be regarded as a property of the system rather than that of a single compound.

The experimental definition is associated with the following behaviour of the chemostat: given a steady state of the culture, changing the feed concentration of S_{lim} (hence the influx $V_{S_{lim}}^{feed}$) while keeping all other parameters constant (including the dilution rate[5] and the feed concentrations of the other nutritients) should eventually lead to a new steady state in which the general state of the cells is the same as in the first stady state. Since this is not a statement easily tackled experimentally, it is usually only required that the biomass composition, the residual concentration of S_{lim} and the associated biomass yield have attained the same values they had in the original steady state [18].

THE *(iv) effluent* can be thought of as an "overflow" of the liquid phase in the main working volume since it is required to be equal to the latter in composition. Importantly, the volume per time unit leaving the chemostat through the effluent is required to be a constant fraction of the main volume specifying the *dilution rate* usually denoted by D [time^{-1}]. The dilution rate should also be held constant during volume changes, which may possibly occur before attaining a steady state. Since gaseous and liquid phases may be interconverted, this condition does not imply that the feed and the effluent fluxes have to be equal.

Note that the above requirement causes the total efflux [mol time^{-1}] of any substance through the effluent to simply to be proportional to its main volume concentration.

Since a chemostat system as defined by (i)-(iv) contains only the quantity D to be set independently and explicitly, the dilution rate is regarded as the main parameter of a chemostat cultivation.

A further independent parameter is the feed concentration of the limiting nutrient S_{lim} which determines the influx $V_{S_{lim}}^{feed}$ for a given dilution rate. However, it follows from the above that this only controls the biomass density, and not the steady state concentrations of external metabolites. This conclusion involves the assumption that biomass density change within a certain range has negligible effects on the cells physiology.

The quantity $1/D$ is called *main residence time* since this is the time which infinitesimally small fluid volumes (which may include cells) would spend on

[5] Note that if the dilution rate - and hence the steady state growth rate - is to remain constant, the volumetric feed influx must be constant too, so $V_{S_{lim}}^{feed}$ can only be controlled by the feed concentration of S_{lim}.

average in the main volume - in an ideal system.

Why does D set the specific growth rate? Balance equations and steady state stability.

To understand why a chemostat cultivation attains a steady state, we now discuss the mass-balance equations.

The amount of any chemical species x in the main volume is potentially increased by influx (if x is provided in the feed), decreased by efflux, and changed by cell consumption or production. The differential equation quantifying the momentarily change of the total amount of x [mol time^{-1}] within the main volume may be written as

$$\frac{d}{dt}C_x^{main}\,V^{main} = \tag{2.1}$$
$$\phi^{feed}\,C_x^{feed} \quad -\phi^{eff}\,C_x^{main} \quad -V^{main}\,C_{bio}^{main}\,v_x^{bio} + v_x^{gas\,in} - v_x^{gas\,ex}$$

where V^{main} denotes the size of the main working volume; C_x^{main} denotes concentration of x in it; ϕ denotes a volumetric flux [volume time^{-1}]; the superscripts $feed$, eff, and $main$ indicate quantities associated with the feed, efflux and main working volume, respectively; C_{bio}^{main} [mass volume^{-1}] denotes the biomass density and v_x^{bio} [mol time^{-1} biomass^{-1}] denotes the rate at which x is consumed (negative if produced) by the cells. The total influx from and evaporation into the gas phase is summed up in the flux terms $v_x^{gas\,in}$ and $v_x^{gas\,ex}$ [mol time^{-1}], respectively.

The temporal change of total biomass, $V^{main}C_{bio}^{main}$ [biomass time^{-1}], itself may be characterised by an analogous formula, however with a few simplifications: we assume that $C_{bio}^{feed} = 0$ (sterile feed[6]), and $v_{bio}^{gas\,in} = v_{bio}^{gas\,ex} = 0$ (no biomass influx or loss through the gas phase). Hence we obtain

$$\frac{d}{dt}C_{bio}^{main}\,V^{main} = -\phi^{eff}\,C_{bio}^{main} + V^{main}\,C_{bio}^{main}\,\mu \tag{2.2}$$

where μ [time^{-1}] denotes the specific growth rate[7] - this could be denoted by $-v_{bio}^{bio}$ in Eq. 2.1, albeit biomass is usually measured as mass, instead of mols of cells.

In addition, from the definition of the dilution rate, we have $\phi^{eff} = V^{main}D$ which simplifies Eq. 2.2 to the standard form

$$\frac{d}{dt}C_{bio}^{main} = C_{bio}^{main}(\mu - D) \tag{2.3}$$

This equation enables us to see that at specific growth rate

$$\mu = D \tag{2.4}$$

i.e. if cells divide once per residence time on average, the biomass is constant.

[6] Which is often nontrivial to achieve, since many bacteria can travel upstream with ease.

[7] A more traditional definition of specific growth rate and its connection to the doubling time $t_d = log2/\mu$ is found in Eq. 2 in [37].

If C_{bio}^{main} is constant, we can solve Eq. 2.1 for the steady state concentration of x, denoted by C_x^{main}, to obtain

$$C_x^{main,sts} = \frac{1}{\phi^{eff}}(\phi^{feed}\, C_x^{feed} - V^{main}\, C_{bio}^{main}\, v_x^{bio} + v_x^{gas\,in} - v_x^{gas\,ex}) \quad (2.5)$$

Hence, steady state concentration of a product ($C_x^{feed} = 0$) changes with the total biomass $V^{main}\, C_{bio}^{main}$ if its specific production rate is constant. This implies that different feed concentrations of the limiting nutrient will lead to steady states with different product concentrations. Hence the definition of "limiting nutrient" implicitly employs the assumption that product concentration changes over a certain range are irrelevant for the biological state of the cells.

Stability and uniqueness of the steady state at $\mu = D$ is implied by the requirement of strictly monotonous dependence of the growth rate μ on nutrient concentrations S_{lim}.

The argumentation runs as follows: let us assume that $\mu < D$. It follows from Eq. 2.3 that this causes a decline in the biomass density C_{bio}^{main} which, via Eq. 2.1, causes the limiting nutrient concentration to rise. Now, the monotonicity condition implies that the specific growth rate μ will increase. Conversely, in case $\mu > D$, an analogous argument predicts the decline of the growth rate. Hence, the steady state at $\mu = D$ is stable and unique.

Note that the above sketched dynamics may not take place if the monotonicity condition does not hold. For example $D > \mu_{max}$ implies that the steady state condition Eq. 2.4 cannot be fulfilled since the growth rate will not increase above μ_{max} with higher nutrient concentration, causing the cells to eventually wash out, resulting in a steady state of limited biological interest. A more exotic example is the case of exceedingly high concentration of S_{lim} such that an increase results in a decrease of growth rate, e.g. due to osmotic stress. In this case, the above theoretical framework predicts the steady state $\mu = D$ (in case it exists) to be unstable, and hence hard to observe in a chemostat[8].

Stability of a fixed point does not exclude oscillations of the system around this point. Oscillatory behaviour may indeed occur in chemostat cultivations - both as an annoyance and as a feature to be studied. Synchronised oscillation in chemostat cultivations is often exhibited by many *S. cereviciae* strains due to cell cycle synchronisation, and must be addressed if undesirable. Strains of the CEN.PK family, on which data for this work is based, are reported to be less prone to cell cycle synchronisation [18].

The mass-balance equations above describe (through the consumption and production rates v_x^{bio}) how the biomass influences its environment. However, in order to complete the description, one would need to answer the reverse question: how is the biomass (more specifically v_x^{bio} and μ) influenced by its environment?

An important qualitative aspect of the answer was already introduced as the monotonicity assumption for growth rate. This information already allowed to analyse steady states. Moreover, since the assumption is a necessary condition for the prediction of the - experimentally observed - stability of steady states, it can be regarded as an experimental fact within the range of observation of

[8] Regulated dilution rate would presumably make observation of such steady states possible.

such steady states. A second qualitative aspect is employed by the concept of "limiting nutrient": the assumption that as long as the concentrations of S_{lim} is the same in two steady states, cells in these cultures are (practically) identical, independently from the other concentrations.

Ideally, a complete, quantitative answer to the above question would allow to predict the dynamic behaviour of the chemostat culture from any given initial state. However, this would require complete knowledge of the relevant biology. Hence there are only partial answers available, typically in the form of approximate, phenomenological formulas.

In many cases the assumption is employed that the growth rate μ generally (not restricted to constant growth rates) only depends on the concentration of the limiting nutrient. Often the Monod-equation[9]

$$\mu = \mu_{max} \frac{S_{lim}}{S_{lim} + K} \tag{2.6}$$

is used to quantify this dependence, where μ_{max} and K are phenomenological constants depending on the experimental conditions. Note that this equation fulfils the monotonicity requirement for the growth rate, and exhibits an asymptotic maximal growth rate, μ_{max}. Naturally, such a phenomenological description has a limited range of validity, for example the prediction to associate any near-zero substrate concentrations with non-zero growth rates is of no direct biological meaning.

Equation 2.6 also allows to estimate the dependence of the steady state residual concentration of the limiting nutrient, which we denote with S_{lim}^{stst}. The resulting expression is

$$S_{lim}^{stst} = \frac{K}{\mu_{max}/D - 1} \tag{2.7}$$

Again, predictions for S_{lim}^{stst} with D/μ_{max} near zero or near one are to be treated with caution. Nevertheless, the general trend is intuitive: the residual steady state concentration S_{lim}^{stst} steeply rises as the dilution rate comes close to the maximal growth rate of the microorganism under the given conditions. This was indeed reported in [96] (s. Fig. 6).

The above drawn picture is of course simplified: in real life, experimental (and financial) limitations are present. Hence, in practice it is usual to monitor only biomass density, as well as concentrations and external fluxes of a few selected chemical species. A practical definition of steady state is to regard at least five main residence times as necessary to reach steady state, which is assumed to have been reached if macroscopically monitored quantities change less than 2 % whithin the next main residence time [18].

Comparing cell cultures from chemostat and from batch cultivations should be undertaken with great care, since the two methods present different microenvironments to the cell. In a chemostat, cells experience a continuous growth control by limitation of a single nutrient, while in typical batch cultures no initial nutrient limitation is present, usually resulting in higher growth rates during the exponential phase of the culture. There may be scenarios during the lag phase of a batch culture when nutrient concentration has diminished to be comparable

[9] In order to increase readability, concentration of the limiting nutrient will be denoted simply by S_{lim} (instead of $C_{S_{lim}}^{main}$ or $[S_{lim}]$) in all formulas.

to that in a chemostat culture, however, this is not necessarily the case for other metabolites (e.g. ethanol) at the same time, which may result in a significantly different environment.

But even if there are time points where all the main parameters - concentrations and growth rate - are comparable between a chemostat and a batch culture, the states of the cells in the two cultures should not be expected to be identical for reasons, to be discussed in the next section.

2.2.2 Selection pressure and data reproducibility

Due to growth over an extended time period under nearly identical conditions, selection pressure in a controlled way is an inherent property of chemostat cultivations, as we will discuss below. Two properties of such cultures will be put into relationship here: high reproducibility of chemostat data and low concentration of the limiting nutrient.

A cell culture is not a copy of perfectly identical cells. Genetically identical cells can exist in various states while adapting to various environments - for example transcriptional activities and enzyme amounts can change, leading to different reactions of the cell to environmental conditions. In this section it is assumed that, in the absence of strong selection, a certain range of states - and corresponding behaviours - is present in a yeast population of genetically identical cells.[10]

On the other hand, as argued below, presence of selection pressure results in a more homogeneous and reproducible distribution of states in the population.

To capture this notion, let us assume that a cell is capable to exist in different *metabolic states* which can - in principle - be distinguished from each other by observations of the cell‘s composition and its microenvironment. For the purpose of studying metabolism, cells exhibiting the same metabolic fluxes and growth rates in a range of microenvironments will be regarded as being in the same state.

It is assumed that the only dynamical quantities directly influencing a cell's state are the local concentrations of substances consumed or produced during metabolism. This means that further parameters (temperature, PH value) can either be fixed in the experiment, or that they do not directly influence the cell state. In particular it is assumed that biomass density does not directly influence the cells[11].

In order to produce reproducible quantitative data, reproducibility of the distribution of cell states in the cultivation is desirable. Clearly, this is a stricter requirement than what is needed to produce qualitative data, in which case it suffices to produce a significant *difference* between populations to be compared. Based on the above assumptions, chemostat cultivations are expected to be more homogeneous and more reproducible with respect to cell state distribution.

Since, in contrast to batch cultures, chemostat cultivations can exist over extended period of time, it is to be expected that the resulting population is more reproducible and more homogeneous with regard to cell state. Moreover,

[10] The cell cycle introduces a necessary inhomogeneity, but this can be averaged out either by using non-synchronised cell populations or by measuring multiple time points, hence this issue will not be addressed.

[11] Even though yeast cells may, in certain scenarios, show quorum sensing, [82].

the assumption of the monotonicity of the growth rate with regard to nutrient concentration implies:

A chemostat cultivation with dilution rate D selects for cells which are able to exhibit a growth rate $\mu = D$ under the lowest concentration of the limiting nutrient.

The argument can be summarised as follows: let us consider two cell populations P_1 and P_2, both completely homogeneous in terms of cell states, exhibiting distinct growth rates $G_1(S_{lim}) < G_2(S_{lim})$, respectively, for a given limiting nutrient concentration S_{lim}. The monotonicity of $G_2(S_{lim})$ implies that there exists some concentration $S'_{lim} < S_{lim}$ such that $G_1(S_{lim}) = G_2(S'_{lim})$. In other words, the population P_2 can achieve the lower growth rate of P_1 already at a lower nutrient concentration S'_{lim}.

Now, let us consider the case that P_1 and P_2 are sub-populations in a chemostat nutrient concentration S_{lim}, and dilution rate $D = G_1(S_{lim})$. The higher growth rate of population P_2 will cause it to gain biomass until the nutrient concentration is reduced to S'_{lim}. Depending on whether or not the population P_1 can grow with the dilution rate D under this concentration, it will or will not remain in the cultivation (in the latter case the cells may shift to a different cell state or will wash out).

Thus, given a chemostat cultivation with a mixture of cell states, the residual nutrient concentration will, after some time, reach the lowest value at which accessible cell states are able exhibit the growth rate D. Eventually only those cell states will remain in the chemostat which are capable of the growth rate $\mu = D$ at this residual concentration which is the statement above.

Since higher growth rate generally requires higher nutrition influx, an immediate consequence of the above is that, all other conditions kept constant, residual nutrient concentration is higher in chemostat cultures with higher dilution rate; we expect the concentration to rise to high levels as the dilution rate approaches the maximal growth rate (c.f. Eq. 2.6). Indeed, this behaviour is well know in sugar limited microbial chemostat cultivations [51, Fig.2], [96].

So far, the genetics of a population was not examined, i.e. accessible cell states are assumed to be limited to those of a given organism of given genotype. In reality however, mutations may eventually start to play a major role due to the prolonged adaptation pressure. On one hand, this opens a fascinating method to study a controlled evolution process, already known to Novick and Szilárd [61], even leading to attempts of developing technologically useful strains [64]. On the other hand, mutations sets an upper limit for exploring the original genotype, hence it is recommended that chemostat experiments aiming at this do not involve more than 20 generations [18].

We now return to the comparison between chemostat and batch cultures. The above considerations point out a further aspect, regarding to which cell populations from the two differ from each other: they are products of different selection pressures. Moreover, based on the above argument, we now have a reason to expect chemostat cultures to produce more reproducible distribution of cell states in the populations.

Applying the above outlined view, a batch culture, too, can be seen as a system selecting for certain cell states. Unlike a steady state chemostat cultivation, it selects for cells exhibiting the highest growth rate, however, during

a time period with constantly changing nutrient and product concentrations. Hence the resulting distribution of cell states is the result of a more complex, less controlled process, and it is realistic to assume that, for each sample, it may depend on the history of the system until when that sample was taken (including its initial condition). Additionally, the selection process in a batch culture is less strict: slower growing cell subpopulations are not washed out as in a chemostat. This effect is weak if, as typically the case, the cultivation period is short compared to a chemostat cultivation where the initial batch-mode is followed by a much longer continuous mode.

To sum up, in cases where the above assumptions are met to a reasonable extent, the distribution of cell states in a batch culture is not only expected to change in time, and converge slower to a final distribution, but also to be much more dependent on the history of the system than in a chemostat cultivation. In accordance to the above, chemostat cultivations, were found a more reliable - but also laborious - source of data than shake flask cultures [63].

2.3 Cell culture conditions

Based on the articles [47] and [97], the following sections contains a brief description of cultivation conditions and extraction of samples, the latter being the basis for generating the datasets described in the next sections.

The yeast strain was CEN.PK113-1A (MATα, URA3, HIS3, LEU2, TRP1, MAL2-8c, SUC2), provided by Dr P. Kötter, Institut für Mikrobiologie, J.W. Goethe Universität, Frankfurt, Germany [46] and stored in 30% v/v glycerol at -80 C.

Cell samples were taken from steady state chemostat cultures of *S. cerevisiae*. Cultivations were set up in 0.8-1 l working volume in Biostat CT bioreactors (of maximal 2.5 l working volume).

The feed medium entering the working volume was minimal medium [94] with 10 g/l glucose as carbon source, containing 10 mg/l ergosterol, 420 mg/l Tween 80 (source of oleic acid), and 0.5 ml/l BDH silicone antifoam.

The chemostat cultures were inoculated to an initial biomass density of 0.5 at $OD_{600\ nm}$, and maintained as batch cultures for 6 - 9 hours, when continuous medium feed was started while the cells were still growing exponentially.

The parameter under study was oxygen availability. Therefore, cultivations with five different oxygenation conditions ranging from aerobic to fully anaerobic were performed. In these, 20.9%, 2.8%, 1.0%, 0.5% or 0.0% oxygen was present in the chemostat inlet gas. In order to keep differences in the culture conditions as small as possible, the total inflow of gas was kept the same in all cultures and lower oxygenation was achieved by replacing oxygen by the equivalent volume of N_2 in the inlet gas.

Cultures supplied with 2.8% or 20.9% oxygen were subject to oscillations. To prevent these, at the time when continuous medium feed was started, ca. 5% of the biomass in the bioreactor was added to the culture as cells in mid-exponential to late exponential phase [101].

In steady state operation the following conditions were maintained: $D = (0.10 \pm 0.02)/h$ (defining the average growth rate in the chemostat), working volume between 0.8 and 1 l, at 30 C. The total gas inflow was 1.5 [volume gas] [volume culture]$^{-1}$ min^{-1}, and pH was kept at 5.0.

Steady-state samples were taken after the cultures had been in constant conditions for a minimum of four main residence times (six generations). The cell samples were transferred to 60% methanol at -40 C immediately after their removal from the bioreactor, and collected by centrifugation.

Steady states were assessed over four to nine residence times (6 to 13 generations) for constant biomass production, carbon dioxide production, oxygen uptake rate, alkali utilization, and extracellular metabolites. The concentration of gases (CO_2, $^{13}CO_2$, O_2, N_2, Ar) was analyzed continuously in an Omnistar quadrupole mass spectrometer (Balzers AG, Liechtenstein), calibrated with 3% CO_2 in Ar.

Most measured quantities were normalised to the dry weight (DW) of the biomass of the sample. For dry weight determination, cells were washed with one to two sample volumes of distilled water, then dried to a constant weight at 100 C. For this measurement, duplicate (5 ml) or triplicate (2 ml) samples were used.

2.4 Quantifying metabolic fluxes by MFA and isotope labelling

Internal flux distribution during steady state growth in the above described cultures was obtained by constrained flux balance analysis. In this approach, metabolic flux balance analysis (FBA) techniques are used together with constraints from external flux measurements and from estimates of intracellular metabolic branching ratios obtained by isotope labelling technique. While using information from the latter dataset opens possibilities for finer analysis, it is connected with high costs and can only be performed on steady state cultures.

Application of FBA on a stoichiometric model of yeast central carbon metabolism [56] resulted in a number of linear equations. Each solution to this equations represents a flux distribution consistent with the stoichiometric model. A number of external fluxes (glucose, ethanol, glycerol, CO_2) were measured. Setting these rates to their experimentally measured quantities results in a reduction of the solution space. Further experimental input came from the estimation of the depletion rate of biomass precursor metabolites which was determined using from known biomass composition and growth rate based on [32].

It should be noted that biomass composition of *S. cerevisiae* was assumed to be the same in all oxygenation conditions studied, since the biomass composition in the two extreme conditions, i.e. in fully aerobic and in anaerobic, had been demonstrated to be essentially the same.

However, this information does not suffice to provide a unique solution for the flux distribution of the central carbon metabolism. In order to obtain further constrains by experimental means, ^{13}C labelling technique and metabolic flux ratio (METAFoR) analysis were used. This approach uses labelling of glucose molecules by the heavy carbon isotope ^{13}C to determine the flux ratios in certain nodes of the metabolic network. The resulting six linear equations were used as additional constraints to solve the underdetermined system of the flux balance equations. For an introduction on the subject see for example Chapter 9 of [85]. Since all measured quantities have experimental errors, the internal flux

distribution was obtained by minimisation of the sum of the weighted square residuals while requiring that the flux balance equations be satisfied exactly.

The scheme of the stoichiometric model used in this analysis is shown in Fig. 2.2 on the left. The glyoxylate cycle was omitted from the model since the METAFoR data showed that the pathway was inactive. The transport of AcCoA, the final step of the cytosolic PDH bypass, was also omitted since exogenous carnitine is required for carnitine shuttle activity, and it was not provided in the medium. A simplification was adopted regarding the pyruvate dehydrogenase (PDH) bypass, depicted in Fig. 2.2 as branching off from the cytosolic acetate producing branch. This enzyme could be partially located in mitochondria and hence contribute directly to the formation of $AcCoA_{mit}$. However, the 13C-labelling technique used cannot reveal the possible contribution of PDH bypass pathway to the carbon flux in mitochondria. As an alternative, expression of the gene ACS1 was analysed and found negligible. Since this gene encodes the mitochondrial AcCoA synthetase which is essential for the contribution of mitochondrial PDH bypass to the formation of $AcCoA_{mit}$, the mitochondrial PDH bypass was not included in the stoichiometric model. The result of the flux analysis is shown in Fig. 2.2 on the right.

In the original publication [47, Jouhten *et al.*], cofactor mass balances were not included in the stoichiometric model. However, the intention of this work is to use the internal flux data in combination with a kinetic model described in Chapter 4 in this thesis. This model contains not only the redox cofactor pair NADH/NAD and its turnover, but also the oxygen uptake rate, as well as a representation of cell respiration which couples these to each other. This made it necessary, to partly repeat the above described constrained optimisation process using a slightly different set of constrains. This calculation will be explained in Section 4.2.3 of this work.

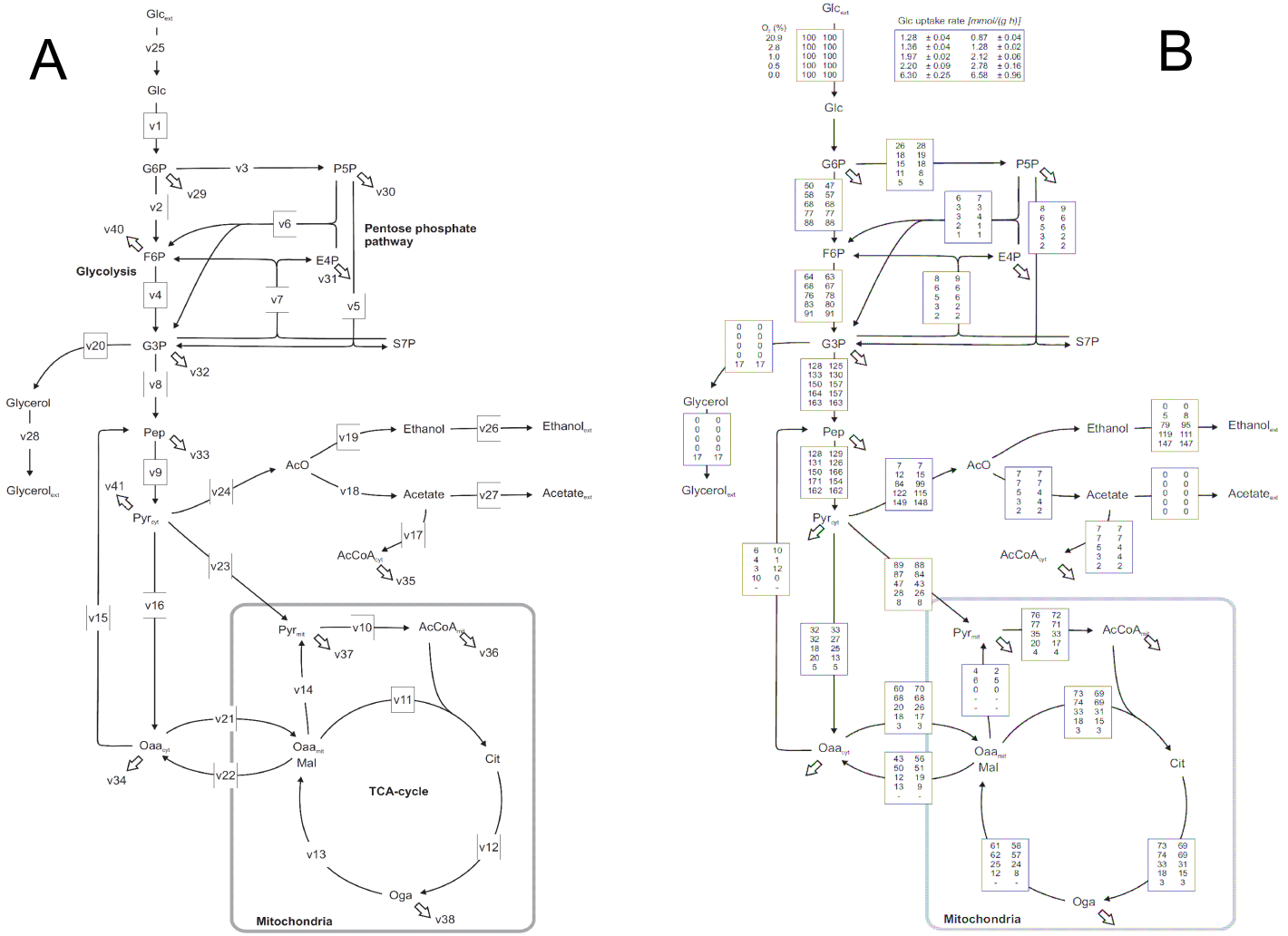

Figure 2.2: *Panel A*: Stoichiometric network model of yeast central carbon metabolism used in [47]. Arrows define positive directions. Subscript *ext* denotes compounds outside the cell. Small arrows denote anabolic fluxes to biosynthetic pathways; their values are defined by the flux balance condition.
Panel B: Result of ^{13}C constrained FBA for *S. cerevisiae* in glucose limited chemostats, subjected to 20.9%, 2.8%, 1.0%, 0.5% and 0.0% oxygen in the chemostat inlet gas. Numbers show relative net fluxes normalised to specific glucose uptake rate (shown at the top) in the corresponding experiment. Oxygenation conditions correspond to rows, as indicated next to glucose intake. Replicate cultures correspond to columns. Based on figures in [47].

2.5 Quantifying gene expression by microarrays

Micro-arrays, as a hybridisation-based measurement technology for mRNA abundance, have increasingly been used from the 1990's on, and have since become standard for genome wide transcriptome studies for organisms with fully sequenced genomes [18].

There are various micro-array technologies using different approaches. An important attribute is whether the micro-array consists of full length cDNAs as hybridisation spots, or of only a few shorter oligonucleotides per ORF. To date, the oligonucleotide-based approach has become standard in quantitative transciptomics due to higher sensitivity and specificity.

With regard to another important classification, micro-arrays can be 'double-dye' or 'single dye'. Double-dye arrays are designed to compare two different biological mRNA samples dyed with green and red dye, respectively, then mixed and hybridised on the same array. The relative intensities of the two colours at the hybridisation site of a certain gene (or oligonucleotide) is a measure of the relative difference in transcript abundances between the two samples. While this approach circumvents serious normalisation issues in a cost-effective way by comparing results from the same hybridisation process, it complicates comparisons of more than two biological conditions, and it does not result in any reliable absolute 'expression value' for later reference.

Single-dye oligonucleotide cDNA micro-arrays have been developed and marketed since the early 2000's by Affymetrix (Affymetrix Genechips). Here, normalisation and quantification issues are addressed by using 10 to 20 gene-specific 25-mer oligonucleotides micromanufactured in high numbers, and distributed over the chip to reduce spatial effects from imperfect mixing during the hybridisation process. To better quantify the specificity of binding of the mRNAs to these oligonucleotides, each hybridisation value from a probe site of copies of a certain oligonucleotide ('perfect match' value) is compared to that from a corresponding probe site consisting of oligonucleotides of the same sequence except for a single-point mutation ('mismatch' value). Since a mutation strongly reduces specific binding while leaving non-specific binding probabilities nearly unchanged, it provides a good internal control for background noise. The dataset from the 'perfect-match' and 'mismatch' oligonucleotide variants are then further processed by statistical and normalisation algorithms which provide a single 'hybridisation intensity' value for each gene.

For some further aspects of oligonucleotide micro-array analysis, the reader is referred to [15] and [98].

If the experimental procedures are otherwise well designed, the reproducibility of these values is expected to be high enough for comparing different biological samples hybridised on different chips. Of course, reproducibility of the biological sample itself is often an issue. In accordance with our previous discussion, chemostat cultivations were confirmed to be more reliable than typical shake flask cultures [18]. It is often assumed that, in typical genechip experiments, a fold difference of approximately two is the threshold to statistically sound 'difference call'. The traditional presentation of the results on a $\log_2$ scale helps identification of genes satisfying this requirement.

Provided that strains, cultivation conditions and protocols are rigorously standardised and observed, chemostat-based micro-array data, even based on experiments from different laboratories, are assumed sufficiently reproducible

to be used in databases containing expression pattern information, making the resulting datasets valuable assets for the systems biology community.

The here described data were produced using chips manufactured by Affymetrix. The results of the experiments were presented by Rintala *et al* in [69].

Affymetrix microarray analysis on samples originating from two (for experiments with 0.5%, 2.8% oxygen) or four (0%, 1.0%, 20.9% oxygen) different cultivations with identical conditions was performed. From cultures with 0.5% and 2.8% oxygen, two parallel steady state samples were also analysed. In addition, four parallel samples from one of the cultivations with 1.0% oxygen were analysed as well.

Each sample was hybridised to the GeneChip Yeast Genome 2.0 Array at +45C during 16 h, according to Affymetrix' GeneChip Expression Analysis Technical Manual.

Data analysis was performed using the software R/Bioconductor version 2.5.1. Raw data was normalised with Robust Multichip Average (RMA) normalisation.

For the purposes of the analysis presented in this thesis, the $\log_2$ normalisation of the result presented in [69], was reversed, in order to obtain the measured fold differences of mRNA in the samples.

2.6 Quantifying intermediate metabolites by HPLC-MS

Part of the obtained samples was used to determine the amounts of certain intermediate metabolites in the central carbon metabolism. The list of these metabolites and the results for the five steady state conditions are shown in Fig. 2.3.

Cell samples used for this purpose were, immediately after their removal from the bioreactor, transferred to 60% v/v methanol at -40 C to quench metabolic processes. Cells were collected by centrifugation at 2000 g at -19 C for 5 min, washed once with 60% v/v -40 C methanol at -19 C [19], then frozen in liquid N_2, and stored at -80 C.

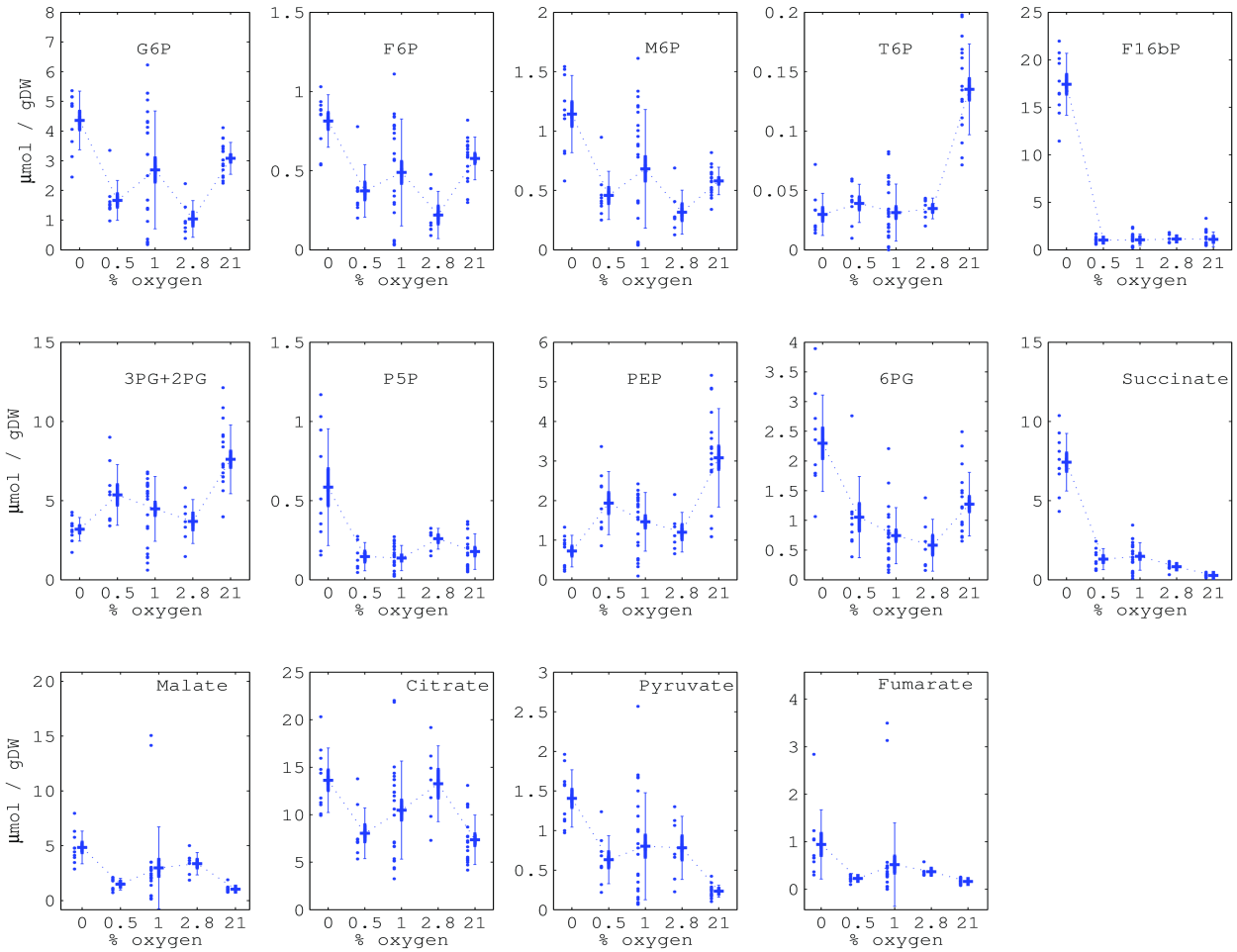

Figure 2.3: Intracellular metabolite levels quantified by HPLC-MS at five different oxygenation conditions. Each point corresponds to a sample. For each condition, 2-4 cultivations were performed, resulting in 7-24 samples. Mean value, standard deviation (thin error bars) and standard error of the mean (thick error bars, used in [97]) were calculated from all samples corresponding to a condition. gDW: cell gram dry weight.

Metabolites were subsequently extracted in boiling ethanol [33], and separated by liquid chromatography (LC) [16] using a Waters HT-Alliance liquid chromatograph. The separated samples were then quantified by a Micromass Quattro Micro triple quadrupole mass spectrometer (MS).

Adenosine nucleotides (ATP, ADP, and AMP) were separated and quantified using ion-pairing LC-electrospray ionization (ESI)-MS with diisopropyl-amine as the ion-pairing reagent. HPLC was carried out in an Agilent 1100 (Santa Clara, CA) with an Xterra MS C_{18} (1 x 150 mm) column (Waters, Milford MA). The nucleotides were detected by ESI-MS (positive ionization mode) of the mass spectrometer. For further details refer to [97].

2.7 Discussion

Regarding experimental information, the kinetic modeller's dream is a dataset which describes the state of the reaction system under study in terms of basic (preferably physical) quantities, such as rates, concentrations and kinetic laws.

The combination of the data described in this chapter represents a considerable amount of information on the cell's central carbon metabolism and comes closer to this ideal than it was possible just a few years ago. Although many datasets are available on central carbon metabolism of yeast, most of these studies puts emphasis on one or two of the following three aspects: monitoring change of enzyme activities (e.g. [9]), quantifying intracellular metabolites [87], or determining intracellular flux distribution via isotope labelling[12] [6]. The datased used in this work represents a remarkable balance of these three aspects, hence a good basis for system biology approach.

Data on intracellular metabolite fluxes and concentrations reveal information on the momentarily steady state of the chemical reaction system, while measuring transcription activities reveals some information on the enzymes' role in achieving this state, as demonstrated for the enzyme fumarase in Fig. 2.4. Naturally, a number of difficulties are present.

Estimating intracellular metabolite concentration from measurement data. The theory of chemical reaction kinetics employs concentration [mol volume^{-1}] of the reactants as a basic quantity. Since the output of HPLC-MS experiments is the absolute amount of a given substance present in the sample, knowledge of the relevant volume of the sample would be necessary. The presented results are, as customary, normalised to cell dry weight of the sample. Hence, the information missing is the value of the combined volume of the relevant organelle of all cells per dry weight of the cell sample. For many purposes, it would be sufficient to know that this quantity is approximately constant across experimental conditions (in our case oxygen levels). However, we were not aware of direct measurement data for these. It is known that the relative size of organelles, as well as the average size of yeast cells may vary between different conditions.
Nevertheless, normalisation to dry weight is assumed by experimenters to produce more robust results than normalisation to optical density (OD) which is often preferred due to the simplicity of the measurement procedure.

[12] Note that although flux estimation via FBA is used in many studies, this is mostly based on external flux measurements only with no constraints from isotope labelling.

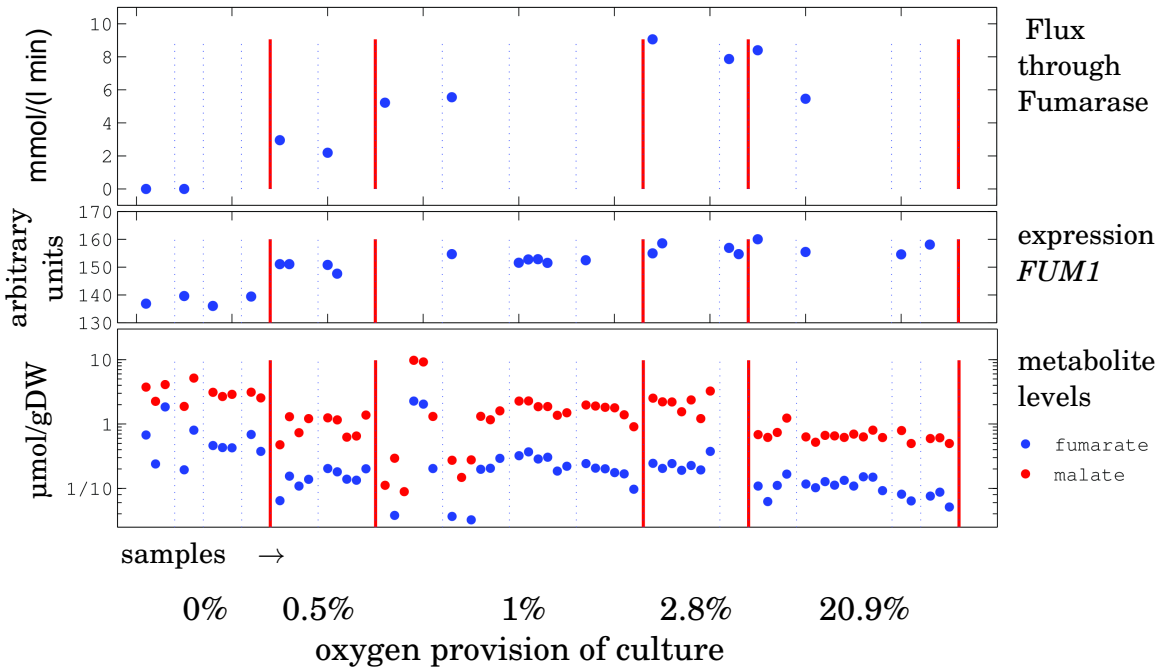

Figure 2.4: The three datasets demonstrated through species associated with the reaction catalysed by Fumarase. Shown are sample data for reactant levels (lower panel, log values), expression levels of the corresponding gene (middle panel), and flux through the reaction (upper panel) from 16 glucose limited steady state chemostat cultivations. Values for each sample are shown, dotted blue lines separate samples from different cultivations, thick red lines separate cultivations with different oxygen levels.
Within the metabolite and the transcription datasets, measurement for different species is performed on one sample. Corresponding values are represented by dots at the same horizontal position (shown for metabolite levels). Values in different panels (and datasets) always result from different samples.
The metabolite level dataset exhibits high noise, while values from the same samples clearly correlate, as highlighted by the logarithmic plot.

In this work, in accordance with the experimental collaborators, 2 ml cytosolic volume per gram dry weight of cell sample is assumed for all conditions. This is consistent with [87], and close to the value implied by 2.38 ml cellular volume per g dry weight as reported in [34] for batch culture.

Glucose can be regarded as the starting point of intracellular metabolism, hence, knowledge of its concentration is desirable in the present context. Extracellular glucose concentration is known to be low (in the 0.1 mM range, [90]) in glucose limited chemostat cultures. Indeed, concentration of residual glucose was for all cultures below the detection limit (see [97] for methods). Measuring intracellular concentration may require higher effort, furthermore, internal level must be lower than the external level, since *S. cerevisiae* uses a (facilitated) diffusion intake mechanism [72], which requires the presence of a

concentration-gradient.

Selection pressure in chemostat cultures has been a known aspect from the beginning [61]. It has been used to study artificial evolution [61], [18], including attempts to achieve technologically advantageous changes [64],[14]. However, in spite of the existence of excellent studies including theoretical aspects, such as [37], we are not aware of an analysis focusing on the characterisation of the selection pressure - e.g. by specifying an objective function - present in chemostats or similar continuous cultivations. We deduced in Section 2.2.2 that - under certain assumptions - chemostat cultivations select for cells which are able to produce the growth rate equivalent to the set dilution rate with the lowest extracellular glucose concentration.

An approach where a precise characterisation of selection pressure in the cultivation is highly relevant, is the prediction of metabolic fluxes via flux balance analysis combined with an objective function to be maximised (or minimised) [42]. Selection pressure in a cultivation naturally influences the metabolic state of the cells, hence, objective functions reflecting this information may be expected to perform better in predicting flux distributions. This is in line with a recent study [78] in which various objective functions in flux balance analysis were tested with regard to their ability to reproduce flux distribution in the metabolic network of *E. coli*. In cells from chemostat cultures, the objective function based on the maximisation of ATP yield per flux unit was the most successful [78, Table II, Fig. 5].

Selection pressure towards lower concentration of the limiting nutrient in continuous cultures implies, for example, competition between species with differing glucose affinity, with the inevitable outcome that the one with the lower affinity is 'outgrown', as the nutrient concentration sinks below its tolerated level for the set growth rate, as reported in [65] (see also Fig. 13 in [96]).

Certain parameters, such as temperature, effect the cell's growth performance. One implication of this is that residual glucose concentration differs for different temperatures [51, Fig. 2].

Aerobiosis allows higher specific biomass production, mainly because carbon loss towards ethanol production is lower or nil. This means, higher oxygen concentrations enable yeast to produce the steady state growth rate at lower glucose influx (c.f. Fig. 7 in [96]). It seems a highly reasonable assumption that, for lower glucose influx, lower concentration of glucose is sufficient. This implies that residual glucose concentrations should be lower in the cultures with higher oxygen concentration. This may be validated if residual glucose concentration in the leftover samples can be measured using more sensitive methods than the online measurements at the time of the experiment.

The form of selection pressure in a continuous culture clearly depends on the feedback loop involved. Biomass growth control in a chemostat is based on a 'built in' negative feedback: faster growth quickly leads to lower residual glucose concentration. A continuous culture with a different control loop may exhibit selection pressure of different characteristics. For example, so-called auxostat cultivations achieve steady state by an explicit feedback loop which reacts to biomass increase in the main volume by an increase in the dilution rate. This mechanism leads to growth at maximum growth rate of the cell type, and accordingly high residual concentration of the nutrient. Such cultures will

exhibit selection pressure towards higher growth rates, however, we expect the exact form of the characterising objective function to depend on the nature of the biomass signal used in the feedback loop. Naturally, such cultures will contain cells in different metabolic states from those in chemostat cultures. This can be expected to be reflected in the prediction success of the associated objective function when applying flux balance analysis to describe the metabolic state of cells grown in an auxostat culture.

A technologically relevant question is, how to design a continuous culture - specially its feedback loop - to realise a given objective function. Examples for objective functions of current technological relevance are maximal yield on abundant raw materials, such as xylose, or lignocellulose [64]. Regarding the latter material, not digestible by yeast, we hypothesise that a possible approach is to use mixed cultures of yeast and one or more 'pre-digestive' microorganisms, grown in a continuous culture designed to exhibit the desired selection pressure. This would involve analysing the dynamics of mixed microbial colonies under specific forms of selection pressure.

Delay in quenching of metabolic processes as a major source of uncertainty. Removing an aliquot from the chemostat cell culture causes the influx of substrates provided in the feed (most notably glucose and oxygen) to be halted. Concentrations of certain internal metabolites and external glucose concentration may change on a sub-second scale, once external glucose concentration is changed, as demonstrated in [19].

Hence, measuring internal metabolite abundances requires fast and controlled halt of the enzymatic reactions involved, termed quenching. In the presented experiments, quenching was achieved by spraying the samples into cold methanol, after removal from the working volume. It is possible that the transfer time was in some cases long enough to allow a substantial change in certain metabolic concentrations, resulting in some substantial but uncontrolled difference between measurements. If this is case, substantial amount of noise might be generated in this process, if the transfer time is not sufficiently homogeneous across samples or cultivations - this is easily possible since not all cultivations were performed at the same time, place or by the same person.

If the 1% condition is disregarded, distribution of the noise levels in malate and fumarate may be hypothesised to result from this source of noise. Relative noise in the levels of these metabolites is higher in the aerobic condition, when flux through the corresponding reaction is significantly higher, presumably generating a higher sensitivity for delayed quenching.

The above issues may be partially responsible for the high noise in the internal metabolite quantification dataset, as seen in Fig. 2.3. Especially in the 1% oxygen condition, interpretation of the raw data is difficult due to high noise. Moreover the noise is expected to contain correlation between different datasets. For metabolite amounts data, typically all metabolites were quantified from the same sample and the correlation between samples is evident from Fig. 2.3. For all data, similar (although, weaker) effects are expected between cultivations.

Chapter 3

The building blocks: single reaction analysis

3.1 Introduction and summary

Before turning to a model comprising significant part of central carbon metabolism, we devote this chapter to the analysis of a few single reaction steps. This approach is stems from the view that chemical reactions steps catalysed by enzymes are natural building blocks of metabolic pathways.

Therefore, this chapter presents a step in integrative analysis of the introduced datasets on metabolite levels, transcription activities, in combination with enzyme kinetics equations. The latter may be interpreted as the integration of existing knowledge about the involved enzymes - e.g. regarding parameter values - but also the means to test the inclusion of allosteric interactions.

The main question posed in this chapter is whether and to which extent the available information is consistent with the measured flux changes through single metabolic reactions, when the discussed steady states cultures with different oxygenation conditions are compared. Consequently, the presented analysis attempts to illuminate to what extent it is necessary to assume further independently regulated regulatory mechanisms to play significant role in flux regulation.

The method presented here is based on estimating flux changes from data on transcription change and reactant amount via enzyme kinetics, and compare this with measured fluxes. Naturally, since transcription measurements have become increasingly easy during the last decades, relating an enzyme's transcription change with the corresponding flux change has become an common method (e.g. [99], [55], [35]). Transcription - flux correlation may be a satisfactory, and for genome-scale metabolic studies it is often the only (easily) available, experimentally accessible regulatory level for flux redistribution.

While it is generally clear that transcription change is only an estimate for protein level or enzyme activity change, it has often been neglected to discuss, to which extend enzyme activity change is related to actual flux change.

It is clear in the light of chemical reaction kinetics that, in addition to changes associated with the catalysing enzyme, a further contribution to flux change comes from the changes in reactant concentration as well as the rest of

metabolism via various interactions.

This regulatory level is often termed *metabolic regulation*. While the details of this influence depend on the form of the effective enzyme kinetics of the reaction, this mechanism is clearly capable to propagate changes from one point of a pathway to another. Hence, the relative contributions of the above mechanisms to flux change also inform us, to which extend flux changes are 'locally regulated' (e.g. transcription) or 'propagated' from other parts of metabolism.

The idea to classify reactions according to this aspect was proposed by the Westerhoff group in [53]. This approach aims at estimating for a given reaction, to which extend the flux change between two states of a cell (e.g. corresponding to different experimental conditions) results from enzyme capacity change and from concentration change of reactants and further effectors. For a quantitative description, the authors introduced the concept of 'hierarchic control' and 'metabolic control' which they quantified mathematically in the form shown in Eq. 3.3. While this formula carries over the notion of 'fold change' from enzyme regulation to flux change, it has some inherent limitations. In this chapter, we discuss these and present an alternative formula for dissecting the flux change into contributions from metabolic and from hierarchic regulation. In addition, we extend the notion to quantify the discrepancy between experimental values and model calculations which offers a valuable measure of the extent to which the approach can describe data.

Enzyme activity change was inferred from transcription data based on mRNA microarray experiments, described in the last chapter. Transcription of a certain protein, however, does not equal its enzymatic activity, moreover, due to the numerous regulatory levels between the two, stating a generally valid relationship of practical use for modelling purposes seems out of reach at this point. Even in cases when the transcription dataset shows clear correlation with the measured flux change, simply equating enzyme activity change with transcription change does not allow reproduction of the latter. To gain quantitative insight into the relationship between transcription and enzyme activity fold change, we introduce a model of the mRNA production cascade. Under the simplifying assumption of *concerted hierarchical regulation*, this model leads us to a one-parameter formula, Eq. 3.18, which we interpret as an approximation to the ideal case, and hence use its free parameter to fit calculated fluxes to experimental data.

As discussed at the end of the last chapter, a further issue was high noise and correlation in the metabolites dataset. The calculations presented in this chapter are based on values from multiple datasets, in which data are expected and partly demonstrated to show correlations across samples or cultivations, as discussed in the previous chapter. Simply pooling all data from all samples for each experimental condition, and using the corresponding average values and errors for the calculations would have erased this information. Consequently, using the error propagation formula (which is based on the assumption of uncorrelated datasets) would have resulted in larger than necessary errors for the calculated quantities. Instead, in an attempt to use information from the experiments in a more efficient way, the implemented calculation method uses single sample values separately, taking into account source sample and cultivation information.

In contrast to traditional data handling which provides only a few statis-

tical properties (such as mean and variance) of the quantities resulting from calculations with data, the presented approach provides a method to generate statistical samples of these quantities.

3.2 Methods

3.2.1 Dissection of flux change according to regulation

Hierarchic regulation analysis

In the following, it is assumed that the total steady state flux J per cell through an enzymatic reaction can be described, on the basis enzyme kinetics, with a formula

$$J(S) = V^{\max} g(S) \tag{3.1}$$

where S denotes the set of concentrations of the reactant and modifier species for the enzyme in question, $V^{\max}$ denotes the maximal activity of the enzyme per cell.

The goal of this section is to discuss and introduce methods comparing two metabolic states associated with two cell cultures under different experimental conditions. For simplicity, we will assume the cultures to be continuous as it is the case in the chemostat experiments this work uses. Quantities introduced in Eq. 3.1 associated with different cell states will be denoted with indices $i = 1, 2$. $J(S_i)$ and $g(S_i)$ will be abbreviated with J_i g_i, respectively. The sign Δ preceding any quantity will be used to denote the difference between values corresponding to the two states.

Westerhoff *et al.* [53] quantified the flux difference by taking the logarithm of the ratio of the respective fluxes:

$$\begin{aligned} log\left(\frac{J_2}{J_1}\right) &= log(J_2) - log(J_1) = \Delta log(J) \\ &\quad \text{while using the right side of Eq. 3.1} \\ &= log\left(\frac{V_2^{\max} g_2}{V_1^{\max} g_1}\right) \\ &= log(V_2^{\max}) - log(V_1^{\max}) + log(g_2) - log(g_1) \\ &= \Delta log(V^{\max}) + \Delta log(g) \end{aligned} \tag{3.2}$$

In order to normalise the terms in last expression to the total relative change one can divide the bottom line in Eq. 3.2 through the first expression to arrive at

$$\begin{aligned} 1 &= \frac{\Delta log(V^{\max}) + \Delta log(g)}{\Delta log\,(J)} \\ &= \frac{\Delta log(V^{\max})}{\Delta log\,(J)} + \frac{\Delta log(g)}{\Delta log\,(J)} =: \rho_h + \rho_m \end{aligned} \tag{3.3}$$

ρ_h and ρ_m are called the hierarchic and metabolic regulation coefficients, respectively [73]. Note that the normalisation makes these quantities independent of the total relative flux change ensuring that their sum equals unity. Using the logarithm of the flux ratio ensures that the choice of the reference state (exchanging state indices) does not affect the magnitude of the result, only its sign. The above definition for ρ_h and ρ_m offers a way to quantify the relative roles of the biological processes associated with $V^{\max}$ and $g(S)$ in Eq. 3.1.

However, applying the above formula for quantifying hierarchic and metabolic regulation has following limitations.

- If either of the two fluxes is zero, Eq. 3.3 produces infinity or zero as an answer, making, e.g. comparison of anaerobic and aerobic conditions impossible for fluxes which are switched off in either state (such as the reaction catalysed by fumarase).

- If J_1 and J_2 have different signs J_1/J_2 is negative, causing the logarithm to become imaginary. This case, however, occurs for many glycolytic enzymes when comparing glycolytic and glyconeogenetic states of metabolism. Leaving out the logarithm, i.e. simply taking J_1/J_2, would, however introduce a dependence of the magnitude of ρ_h and ρ_m on the choice of the reference state.

- If the targeted flux needs to be regarded as the sum of the flux through two or more enzymes (for N enzymes, $J = \sum_{k=1}^{N} J^k$), it is not immediately clear how ρ_h and ρ_m should be calculated. However, this can be necessary if different isoenzymes catalysing the same reaction are studied, or, if different enzymes are associated with opposite directions of the same reaction, as PFK and FBP in glycolysis.

Mathematically, the first two points are connected to the fact that the function $f(x, y) = x/y$ diverges and switches sign whenever either variable passes through zero. One of the possible ways to tackle the third issue is to use a formula in place of Eq. 3.2 which has the property of being linear in J. The above issues motivated the author to find an alternative mathematical formula for quantification of the role of the various components for flux change to aid the subsequent analysis.

An alternative formula for hierarchic and metabolic regulation

Based on the considerations in the last section, one can argue that the (logarithmic) ratio of two fluxes - quantities with range including zero and negative values - is not a natural choice when quantifying their relationship. An alternative possibility is to focus on the flux difference $\Delta J = J_2 - J_1$ instead of the flux ratio. ΔJ, regarded as a function of J, possesses an important mathematical property: it is linear in J. For example[1], if $J = aJ^A + bJ^B$ with two fluxes J^A,

[1] The actual proof using the general case (for a finite dimensional vector space) $J = \sum_k c_k J^k$ is analogous but less instructive.

J^B, and constants a and b, then

$$\begin{aligned}\Delta(a\,J^A + b\,J^B) &= a\,J_2^A + b\,J_2^B - (a\,J_1^A + b\,J_1^B)\\ &= a\,J_2^A - a\,J_1^A + b\,J_2^B - b\,J_1^B\\ &= a\,\Delta J^A + b\,\Delta J^B \end{aligned} \tag{3.4}$$

The linearity property resolves the third issue mentioned in the previous section and allows to extend the result for sums of fluxes.

In this section, we present a way to write the flux difference as a sum of two terms which can be interpreted to quantify the relative roles of hierarchic and metabolic regulation in achieving this flux difference. In purely mathematical terms, the result can easily be summarized in Equations 3.7 and 3.8 as a discrete form of the total differential of the function $J = J(V^{\max}, g)$. This highlights a property of the general enzyme kinetic formula Eq. 3.1: being the product of $V^{\max}$ and g, J is linear in both, hence their linear changes fully characterise the change in J.

Rather than simply stating the formula, we offer certain a derivation to highlight intuitive aspects and to aid interpretation. First, we write the flux difference in the following form:

$$\begin{aligned}J_2 - J_1 &= \\ \Delta J &= V_2^{\max}\,g_2 - V_1^{\max}\,g_1\\ &= V_2^{\max}\,g_2 + V_1^{\max}(g_2 - g_1 - g_2)\\ &= V_1^{\max}\,(g_2 - g_1) \;+\; (V_2^{\max} - V_1^{\max})\,g_2\\ &= V_1^{\max}\;\Delta g \;+\; \Delta V^{\max}\;g_2 \end{aligned} \tag{3.5}$$

The last line in Eq. 3.5 expresses the fact that the total flux difference through a single reaction can be thought to be achieved by first changing metabolic levels to change g_1 to g_2 while keeping the enzyme activity unchanged ($V_1^{\max}\;\Delta g$), then changing the enzyme activity ($\Delta V^{\max}\;g_2$) while keeping the metabolic levels. However, this order is an arbitrary choice, since this can be done the other way around as well. Indeed, a short calculation analogous to the one shown in Eq. 3.5 results in:

$$\Delta J \;=\; V_2^{\max}\;\Delta g \;+\; \Delta V^{\max}\;g_1 \tag{3.6}$$

expressing the above mentioned alternative choice. In order to avoid this arbitrary decision, we take the average of equations 3.5 and 3.6. Since the left sides are the same, this only effects the right side, leading to a surprisingly elegant formula

$$\begin{aligned}
\Delta J &= \frac{(V_1^{\max} + V_2^{\max})\ \Delta g\ +\ \Delta V^{\max}\ (g_2 + g_1)}{2} \\
&= \frac{(V_1^{\max} + V_2^{\max})\ (g_2 - g_1)\ +\ (V_2^{\max} - V_1^{\max})\ (g_2 + g_1)}{2} \\
&= \text{mean}(V^{\max})\ \Delta g\ +\ \Delta V^{\max}\ \text{mean}(g) \\
&= \Delta J_M \quad + \quad \Delta J_H \qquad (3.7)
\end{aligned}$$

with

$$\begin{aligned}
\Delta J_M &:= \text{mean}(V^{\max})\ \Delta g \qquad \text{metabolic contribution to flux difference} \\
\Delta J_H &:= \Delta V^{\max}\ \text{mean}(g) \qquad \text{hierarchic contribution to flux difference}
\end{aligned}$$

ΔJ_M and ΔJ_H may be used to quantify the role of metabolic and hierarchic regulation.

Analogously to the approach in Eq. 3.3, normalising Eq. 3.7 to the total flux change results in quantities independent of it:

$$\begin{aligned}
1 &= \frac{\text{mean}(V^{\max})\ \Delta g}{\Delta J} \quad + \quad \frac{\Delta V^{\max}\ \text{mean}(g)}{\Delta J} \\
&= R_M \quad + \quad R_H \qquad (3.8)
\end{aligned}$$

where

$$R_M := \frac{\text{mean}(V^{\max})\ \Delta g}{\Delta J} \quad \text{(normalised metabolic contribution to flux difference)}$$

$$R_H := \frac{\Delta V^{\max}\ \text{mean}(g)}{\Delta J} \quad \text{(normalised hierarchic contribution to flux difference)}$$

A comparison of the two mathematical approaches for a number of instructive cases is presented in Table 3.1. It should be noted that both kinds of coefficients (ρ and R) defined above can attain negative values. Since the coefficients for metabolic and hierarchic regulation sum up to one, this implies a value greater than unity for the other one.

This reflects cases in which the two components of the kinetics in Eq. 3.1 change in a way that they act antagonistically. This may occur, for example, if increase of product accumulation is counterbalanced by increase in enzyme amount.

J_1	J_2	$g(S_1)$	$g(S_2)$	$V_1^{\max}$	$V_2^{\max}$	ρ_m	ρ_h	R_M	R_H
flux increase									
metabolic									
1	2	0.1	0.2	10	10	1	0	1	0
mixed									
1	2	0.1	$0.1\sqrt{2}$	10	$10\sqrt{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
antagonistic									
1	2	0.2	0.1	5	20	-1	2	-1.25	2.25
zero reference flux									
0	1	0	0.1	10	10	n.a.	0	1	0
flux reversal									
metabolic									
1/2	-1/2	0.1	-0.1	5	5	n.a.$^{(*)}$	0	1	0
mixed									
1/3	-2/3	0.1	-0.1	$3.\dot{3}$	$6.\dot{6}$	n.a.	n.a.	1	0
net flux constant									
antagonistic						using Eq.3.2		using Eq.3.7	
1	1	0.1	0.2	10	5	0.7	-0.7	0.75	-0.75

Table 3.1: Comparison of the two quantification approaches for metabolic and hierarchic regulation for a number of scenarios. The two flux states compared are denoted by subscript $i = 1, 2$. The reaction kinetics is assumed to obey Eq. 3.1. The columns show total flux J_i, resulting from differences in $V^{\max}$ and metabolic concentrations S and the values of the coefficients ρ and R, calculated from their definitions, Eqs. 3.3 and 3.8, respectively. For the case with no net flux change (last row), the formulas without normalisation were used. The cases are qualitatively classified for easier interpretation according to regulation: purely metabolic, mixed, and antagonistic (ie. when metabolic and hierarchic regulations work against each other). Purely hierarchic cases are not shown, since the calculation is the same as for the corresponding purely metabolically regulated cases.
$^{(*)}$ This specific case leads to $\frac{i\pi}{i\pi} = 1$, however, the result is only real if the equation $g(S_1) = -g(S_2)$ holds exactly.

Flux changes not accounted for by the model

There is always some discrepancy between calculated values based on a model, and the corresponding measured values. In the present context, the flux change calculated from the enzyme kinetics model (Eq. 3.1) on the basis of transcription data will not exactly match the measured flux change. We can easily extend the above introduced formalism to quantify this discrepancy by modifying Eq. 3.7 to

$$\Delta J \quad = \quad \Delta J_M \quad + \quad \Delta J_H \quad + \quad \Delta J_{unk} \tag{3.9}$$

where ΔJ_{unk} denotes the flux difference contribution from 'unknown' mechanisms, the prediction of which was not achieved by the information obtained from the above described datasets. The normalised unknown contribution to flux difference

$$R_{unk} := \frac{\Delta J_{unk}}{\Delta J} \tag{3.10}$$

completes the definitions in Eq. 3.8. As before, from Eq. 3.9 follows

$$R_H \quad + \quad R_M \quad + \quad R_{unk} \quad = \quad 1 \tag{3.11}$$

which implies that at least one of the terms must be positive.

For the following analysis, values for R_{unk} were calculated from R_M and R_H according to this last equation.

3.2.2 From transcription data to enzyme activity

While transcription data do enable us to gain information on an enzyme's activity and change thereof, in terms of quantitative changes, they have only indicative value. This section describes how the transcription dataset was used to estimate enzyme activity changes.

Genome-wide microarrays offer insight into the transcription dynamics of nearly any gene of interest. On the other hand, using such datasets to estimate qualitative fold changes in mRNA numbers has drawbacks. This mainly stems from the complexity of the measurement process, typically involving hybridisation, dying and fluorescence measurement. The statistical algorithms used for data processing at different stages are primarily targeted at the *significance* of the difference between two datasets, rather than to the (harder) task of quantifying this difference. Reported fold changes are also known to be rather dependent on the normalisation used in the analysis. For the above reasons, we do not assume high-throughput microarray data values to exhibit simple linear relationship with the mRNA numbers.

On the biological side, messenger RNA production and presence is only one step - albeit an important one - in an array of processes leading to active protein. Many processes involved in changing protein activity are known, such as translation into polipeptides (by ribosomes), folding (often involving chaperon proteins), mRNA- and protein splicing, and, more recently microRNAs (not in fungi). Most, if not all of these processes are known to be regulated via their respective catalysts. Undoubtedly, a number of processes and their respective regulatory mechanisms have yet remain to be uncovered and unquantified. The below introduced general mechanism leaves the question open, how many steps there are in the full production or regulatory process.

In order to gain understanding of the relationship between transcription and enzyme activity, we aimed to represent regulatory aspects of the *de novo* production process in the model presented below. The result of the below presented, rather qualitative discussion is the statement that, for cases where the model assumptions hold, the fold change in transcription and in enzyme activity exhibit a relationship of the form of Eq. 3.18. The main simplifying assumption in the following derivation is that of 'concerted regulation', as formulated in Eq. 3.16.

For the following we assume that the production of the active form of a protein can be written as the result of a cascade of a number of processes each of which employs the product of the previous one as a catalyst, starting with a substrate X_0. We assume that the production of each intermediate depends linearly on the quantity of the previous one such that the dynamics may be written in the following form

$$\begin{aligned}
\frac{d}{dt}X_1 &= \alpha_1 X_0 - \omega_1 X_1 \\
\frac{d}{dt}X_2 &= \alpha_2 X_1 - \omega_2 X_2 \\
&\vdots \\
\frac{d}{dt}X_n &= \alpha_n X_{n-1} - \omega_n X_n
\end{aligned} \tag{3.12}$$

Each coefficient α_i describes the efficiency of the corresponding process. The lifespan of each intermediate X_i is limited by its decay, expressed as $\omega_i X_i$.

For example, X_1 may be associated with a transcript level produced by the transcription process, which is here assumed to depend linearly on the number of transcription factors, which in turn, may be associated with X_0.

X_n is associated with enzyme activity, the level of which we seek to relate to the level of mRNA.

The (non-trivial) steady state X_i^* of the above system is defined by cancellation of the right sides in Eq. 3.12, which occurs at

$$X_i^* = \frac{\alpha_i}{\omega_i}X_{i-1}^* = \frac{\alpha_i}{\omega_i}...\frac{\alpha_1}{\omega_1}X_0^* \quad \text{for} \quad i = 1...n \tag{3.13}$$

We now analyse the effect of regulation on this cascade as a result of the adaptation of the cell to a different environment. In the above model, the effect of this regulation is assumed to consist of modifying the efficiency of the above processes (quantified by the α_i) and the lifetime of the intermediates (quantified by the ω_i):

$$\begin{aligned}
&X_0 \longrightarrow a_0 X_0 \\
&\alpha_i \longrightarrow a_i \alpha_i \qquad \text{and} \quad \omega_i \longrightarrow o_i \omega_i
\end{aligned} \tag{3.14}$$

Eq. 3.13 gives the following adjustment for steady state levels

$$X_i^* \longrightarrow X_i^{*\text{mod}} = \frac{a_i}{o_i}...\frac{a_1}{o_1}a_0 X_i^* \qquad \text{for} \quad i = 1...n \tag{3.15}$$

The message from this formula is that regulatory events simultaneously occurring at all levels may cause far higher level change for a product near the end of the cascade than near the beginning of the cascade.

In order to quantify this effect, we now introduce a simplification in that we assume that regulation of the constituting processes occurs in a concerted way (for example, as a consequence of a single signal), and each process is effected (approximately) to equal extent. In the present context, this *assumption of concerted regulation* is formulated as

$$a_0 = a_1 = \frac{1}{o_1} = ... = a_n = \frac{1}{o_n} \tag{3.16}$$

Together with Eq. 3.15, this gives for the fold change of an intermediate

$$\frac{X_i^{*\,\mathrm{mod}}}{X_i^*} = a_0^{2i+1} \qquad \text{for} \quad i = 1...n$$

resulting in the following relation for the fold change of transcript level mRNA (which we associate with X_1), with active enzyme PROT (which we associate with X_n)

$$\frac{X_n^{*\,\mathrm{mod}}}{X_n^*} = a_0^{(2n+1)} = (a_0^2)^{(2n+1)/2} = \left(\frac{X_1^{*\,\mathrm{mod}}}{X_1^*}\right)^{n+\frac{1}{2}}$$

or

$$\frac{\mathrm{PROT}^{\mathrm{mod}}}{\mathrm{PROT}} = \left(\frac{\mathrm{mRNA}^{\mathrm{mod}}}{\mathrm{mRNA}}\right)^{n+\frac{1}{2}} \tag{3.17}$$

While the exact form of the above formulas depends on the details of the involved processes, Eq. 3.17 suggest that, for a broad class of models, the relationship of the fold changes of protein and mRNA levels has a strong dependency on the length of the cascade of the involved processes, especially in the case when 'concerted regulation' occurs - which may only partially account for changes in enzyme activity.

Based on Eq. 3.17, in this work we will assume the following relationship between enzyme activity fold change ($\frac{V_i^{\max}}{V_0^{\max}}$) and transcription fold change:

$$\frac{V_i^{\max}}{V_0^{\max}} = \frac{\mathrm{PROT}_i}{\mathrm{PROT}_0} = \left(\frac{\mathrm{mRNA}_i}{\mathrm{mRNA}_0}\right)^{trexp} \tag{3.18}$$

where subscripts denote the respective experimental condition; the reference condition is denoted by subscript 0. Since the catalytic constant $k_{cat} = V^{\max}/\mathrm{PROT}$ is assumed not to change between conditions, it drops out from the above equation.

We will regard the exponent *trexp* as a free parameter. This preserves the qualitative difference between experiments[2], but adjusts their magnitude.

Bearing in mind the above simplifications, we will use Eq. 3.18 to estimate active protein fold change from the fold change of transcription data.

This 'transcription amplification' may not only be interpreted to account for post-transcriptional regulation, but also for the above-mentioned potential nonlinearity of the measurement method. In order not to over-amplify transcription changes, we will accept parameter estimation results for the parameter *trexp* up to an upper limit of 5 which is slightly higher than the currently known number of steps in *de novo* enzyme production.

3.2.3 Sample-wise calculations with data

A customary procedure to treat measurement errors is calculating means and standard deviation of data samples[3], then use the 'error propagation' formula to

[2] To be exact, the logarithm of the ratios of fold changes between experiments is preserved by Eq. 3.18

[3] in this section, the word 'sample' will refer to the statistical term meaning finite sample from a probability distribution (German: Stichprobe). References to 'sample from a cell

quantify the resulting uncertainty in derived quantities, such as Eq. 3.8. This procedure approximates the distribution of the data by a Gaussian (defined by two parameters mean and standard deviation) and constructs the distribution for derived quantities.

The assumption of Gaussian data distribution seems like a reasonable approximation in many cases, however for concentration values it produces standard deviations well into the negative range. Since this is not readily interpretable (e.g since some kinetic equations would produce imaginary values), it hinders correct error propagation treatment. One possibility is to choose a different distribution with non-negative range (such as chi-square distribution) as statistical model. However, while this would lead to a more involved error propagation treatment, it would still be an approximation. As discussed in the last chapter and shown in Fig. 2.4, especially the metabolite dataset contains statistical samples which not only show high correlation with measurements from the same aliquot, but also seem not to be described by Gaussian distribution (c.f. G6P and F6P sample data from the 1% O_2 experiments).

This high, correlated, and apparently non-Gaussian noise in many data samples motivated a more accurate, at the same time conceptually simpler treatment. In our approach, instead of first calculating mean and variance, each single sample value (associated with an aliquot for metabolite and transcription datasets, and with a cultivation in the flux dataset) for measured quantities was directly used in the calculations. The sample values were used in the mathematical calculations to generate new samples of values for the calculated derived quantities. For example reaction rate values were calculated from Eq. 3.1 using each sample value from the expression and metabolic dataset.

The procedure may be summarised as follows:

- Case 1: calculations with species with sample-wise correspondence: statistical samples from two or more species exhibiting the property that the single sample values of different species correspond to one another.

  In the present work, there were two datasets which fell into this category: metabolite and transcription dataset values. In both of these datasets, each aliquot measurement resulted in a value for each species (e.g. each genes) within this dataset. Hence these measurements resulted in the same number of values for each species within one dataset, and measurement values (for two different genes) from the same aliquot may be regarded as corresponding to each other. Metabolite data sample correspondence is shown using the example of fumarate and malate in Fig. 2.4 in which the aforementioned strong correlation of the values from the same aliquot is visible.

- Case 2: calculations with species with no clear sample-wise correspondence. For example a metabolite species and transcription values for a gene are examples of two species which were obtained based on different aliquots (partly stemming from different cultivations), usually (not necessarily) resulting in different number of single-sample values. Hence sample-wise correspondence is not the case. All combinations of species for which sample wise correspondence was not seen fitting, were categorized as Case 2.

culture' (aliquot) will be explicitly named.

Measured sample values were treated as follows:

- (i) If only species with sample-wise correspondence appeared in a formula, the calculation was simply performed sample-wise, resulting in a sample for the derived quantity having the same number of points as the datasets on which the calculation was based. For example the calculation of the derived quantity $g(S, P)$ in Eq. 3.19 (c.f. Eq. 3.1) require (by definition) only metabolite species data. Hence, even if the formula contained many species (reactants and effectors), it was calculated sample wise. Note that this approach *inherently takes into account correlation* or other forms of statistical relationship between sample values of different species, replacing the calculation of the variance-covariance matrix in traditional error propagation treatment.

- (ii) If only species with no sample-wise correspondence appeared in a formula, each possible combination was computed. Let us consider a derived quantity $F(\mu, \gamma)$ requiring concentration values from a metabolite species μ and transcription values of a gene γ, for which single-sample values $\{\mu_1 ... \mu_n\}$ and $\{\gamma_1 ... \gamma_m\}$ are given, respectively. The generated sample for $F(\mu, \gamma)$ is given by the set $\{F(\mu_i, \gamma_j)\}_{i=1...n,\ j=1...m}$.

- (iii) If species both with and without sample-wise correspondence appeared in a formula, only those combinations were computed for which the species with sample wise correspondence had matching sample-values.

  As an example, consider a derived quantity $F(\mu, \nu, \gamma, \delta)$ requiring concentration values for metabolites μ and ν with single-sample values $\{\mu_1 ... \mu_n\}$ and $\{\nu_1 ... \nu_n\}$, as well as transcription values from genes γ and δ with sample values $\{\gamma_1 ... \gamma_m\}$ and $\{\delta_1 ... \delta_m\}$. The generated sample for $F(\mu, \nu, \gamma, \delta)$ is given by the set $\{F(\mu_i, \nu_i, \gamma_j, \delta_j)\}_{i=1...n,\ j=1...m}$, having the same number of single values as the previous example.

- (iv) A derived quantity with a generated sample of values was treated the same as directly measured sample values in subsequent calculations.

A non-trivial example of the application of the above procedure is the calculation of the flux contribution coefficient defined in equation Eq. 3.8 This requires values for the derived quantities $g(S, P)$, and $V^{\max}$, as well as the measured flux J for two different conditions. First, the quantities $g(S, P)$ and $V^{\max}$ were calculated independently for each condition, according to (i). This resulted in samples of values of these quantities for the two conditions, which, however, do not have sample-wise correspondence. These values were combined, according to (iii) and (iv), with the measured flux sample values J to calculate a sample of values for the coefficients defined in Eq. 3.8.

The application of this computation scheme is sketched in Fig. 3.1 showing predicted fluxes computed from concentration and transcription data. As a comparison, the result of the traditional treatment - using the error propagation formula and the variance-covariance-matrix to capture correlations - is shown.

The output of this procedure is samples of values for the calculated quantities. These samples can then be used for further calculations. This can either be in their original form, or, if a simpler procedure is preferred in subsequent steps,

the statistical properties of the samples can be captured via standard methods of descriptive statistics, such as calculating mean and variance.

Consequently, calculation of mean and standard deviation for the resulting samples is a mere presentation technique for the end result, rather than a computational tool.

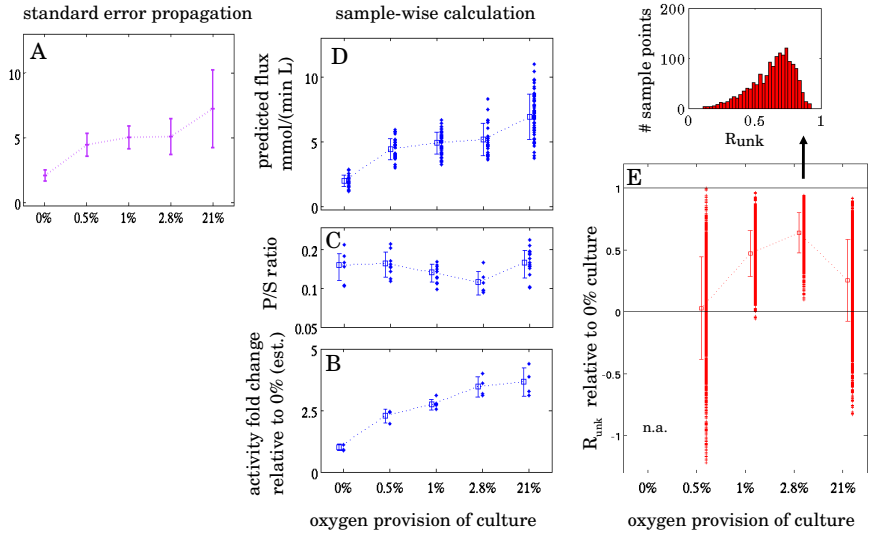

Figure 3.1: Demonstration of sample-wise calculation on the example of the datasets associated with fumarase. Standard error propagation procedure uses variances and covariances of the data samples to estimate standard deviation for the calculated quantity, in this case flux through fumarase (panel A). For non-Gaussian data distributions, this may give limited information on the actual distribution of the result.
In contranst, sample-wise calculation uses all possible combinations of individual data sample values, in this case *FUM1* activity (panel B) and metabolite quantities (panel C, logarithmic scale) to generate a sample for the calculated flux through the enzyme (panel D). For this calculation, data corresponding to different oxygen provision conditions were treated separately.
Coefficients such as R_{unk} compare two conditions involving sample values of calculated and of measured fluxes, hence leading to a high number of single data points (panel E). The resulting distribution of R_{unk} (c.f. Eq. 3.11) for the 2.8% condition is shown as a histogram.
For sample wise calculations, mean and standard deviation of each sample is shown for presentation purposes only, they are not necessary for the calculations.

3.2.4 Fitting to experimental fluxes and parameter estimation

To investigate, to what extent the experimental dataset under discussion is compatible with existing enzyme kinetics knowledge on the scale of single reactions, we compared experimentally determined flux data to reaction rates produced by enzyme reaction kinetics of the form of Eq. 3.1, combined with the appropriate concentration and transcription data.

Since the number of parameters was low, and the non-linearity of a single enzyme kinetics were expected to be of limited extent, we used simple Monte-

Carlo parameter search to answer the question, whether the experimental fluxes were reproducible by a given combination of model and data. This method proved sufficient and it was usually observed that the fit to the flux data was only marginally improved after less than half of the total number of iterations and repeated runs produced essentially the same goodness of fit.

Due to lack of knowledge about absolute magnitudes of V_{max}, and to focus on investigating the *changes* in flux, one of the five experimental conditions was selected to be the reference state (this was the 0% condition with the exception of fumarase). The maximal velocity for the reference state, $V_{max\ 0}$, was then determined by the condition that calculated rate and measured flux match each other.

Prior to the application of the estimation algorithm, for each of the five experimental conditions, outliers in the respective samples S_{cond} for the needed metabolite species were eliminated (c.f. Fig. 2.4). A sample value was regarded as an outlier if it lied outside the double of the standard-deviation range $mean(S_{cond}) \pm 2\, std(S_{cond})$. This resulted in 0 - 4 sample values per species to be excluded.

The following procedure was followed at each iteration of the parameter estimation:

- Random parameters were generated from uniform distributions defined by the limits of the search range for each parameter.
- The value of $V_{max\ 0}$ was calculated from the condition that, for the reference state, the mean of the single sample values for the enzyme kinetics rate equals the mean of the (two) experimental flux data values.
- V_{max} values for the other four conditions were calculated from the transcription fold change of the associated genes according to the formula[4] Eq. 3.18 (depending on the current value of the transcription amplification parameter $trexp$)
- The rate $V_{kin} = V_{kin}(S)$ was calculated from the metabolite data (S) according to the enzyme kinetics expression chosen.
- The discrepancy from the experimental flux J_{exp} was quantified by the following objective function f_o.

$$f_o = \sum_i \left(\frac{J_{\exp\, i} - V_{\mathrm{kin}\, i}}{\sigma_i} \right)^2$$

$$\text{where} \qquad \sigma_i = (J_{\exp\, i}/10 + \epsilon)$$

  The index i refers to combinations of single sample values used in the calculation, as explained in the last section. This approach gives equal weight to calculations resulting from each permitted combination of single measurement values.

[4] In order to save computation time, the calculation of transcription fold changes was not performed using the sample-wise calculation method, rather simply dividing through the mean of the reference state values.

Underlying notion for the form of the data-error term σ is the assumption that the uncertainty of J_{exp} is roughly a fixed percentage of its value. The small value ϵ was introduced in order to limit the influence of J_{exp} values near or equal zero. We chose $\epsilon = 1$ mM min^{-1}, which is in the order of the lowest fluxes in this dataset.

This parameter estimation process produced - besides the parameter values - the best found values for the calculated quantities V_{max} and reaction rates V, for each condition. Based on these quantities, the flux dissection coefficients R_H, R_M were calculated according to Eq. 3.8, and R_{unk} according to Eq. 3.11.

As explained in the last section, these results were initially given as finite samples. For the purpose of presentation, mean and standard distribution for each were calculated. These results are presented in the figures of the next sections.

In order to test the robustness of the results, the above process was repeated a number of times to test uncertainties in the process. This enabled us to detect parameters which were not well defined by the model. This was tested by comparing the standard deviation of the sample resulting from the repeated estimation runs: if a parameter's optimal values exhibited variance comparable to the variance of the uniform distribution - i.e. roughly 1/3 of the parameter's search range - it was marked. Our interpretation in this case is that the estimation process did not identify tighter boundaries than those defined by the search range limits. This may be caused either by near-zero sensitivity of the objective function to this parameter, or by a relationship thereof with one or more other parameters such that various combinations of these result in the same fit.

However, since it was not our goal to estimate kinetic parameter values, this did not alter our conclusions, since all quantities playing a role in the interpretation of the results (which are plotted in the next sections) displayed remarkable stability between repeats. Typically, plots of the results from different runs differed only very slightly, often not easily distinguishable by visual inspection.

3.3 Analysis of single reactions

In the remainder of this section, we apply the above introduced machinery to analyse a few selected enzymes. We aim to investigate, to what extent enzyme kinetics, combined with the given experimental dataset - transcription and concentration - is capable to reproduce experimentally measured flux changes for the natural building blocs of pathways: enzymatic reactions.

3.3.1 Three reactions of central carbon metabolism

Fumarase

We start our analysis with a comparatively simple example, the enzyme fumarase (EC 4.2.1.2). It converts fumarate (fumaric acid) to malate (L-malic acid) in the TCA cycle. In *S. cerevisiae*, the only associated gene is FUM1 (Systematic Name: YPL262W).

Fumarate is potentially activated by several anions, including inorganic phosphate, however, since we do not have any data on these, we did not include them in the model.

Its kinetic properties are far from trivial: For fumarate concentrations up to 1 mM the enzyme is well described by Michaelis-Menten kinetics. At higher concentartions up to 30 mM, fumarate has an activating effect, while for concentrations above 100 mM fumarate inhibits the enzyme [57].

While this presents a puzzle to enzyme kinetics, fumarate concentrations

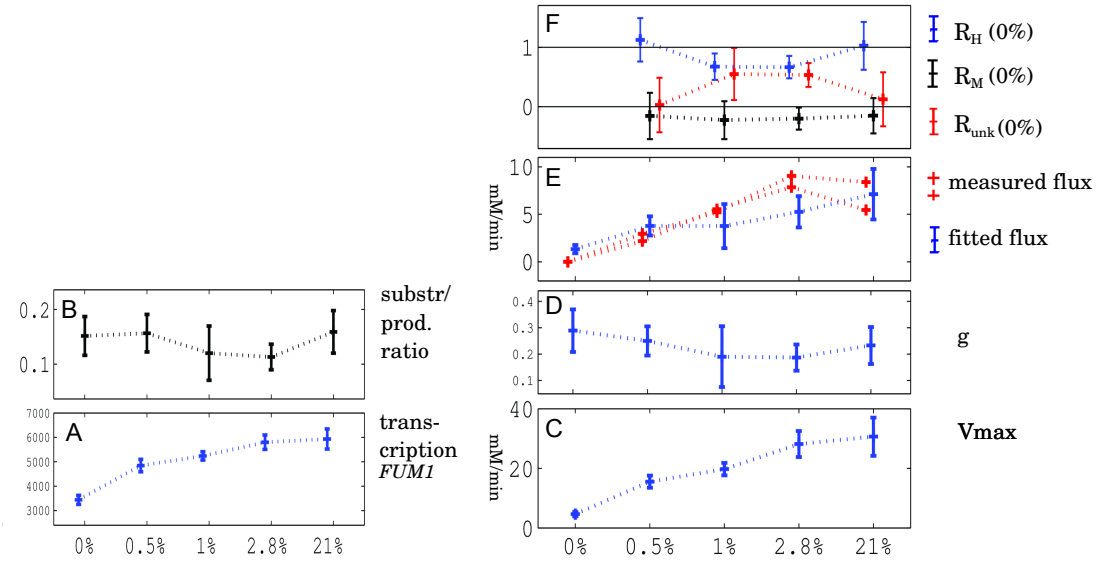

Figure 3.2: Flux dissection analysis for fumarase based on fitting of reaction rate to measured flux data. Transcription and metabolite quantification datasets are indicated in panel A and B, respectively. Results from a parameter search with 800 iterations: Values for Vmax and g are shown in panels C and D, respectively. Corresponding best fit reaction rates and corresponding measured flux data points are shown in panel E. Flux dissection coefficients R_H, R_M, and R_{unk} calculated with respect to the 0% condition (see Eqs. 3.8 and 3.9, note that there is no assigned value for the reference condition) are shown in panel F.
Calculations were based on single sample values, shown are means and standard deviations of the resulting statistical samples, except for experimental flux analysis for which both cultivations are shown separately.

The measurement data and calculated quantities associated with this enzyme are shown in Fig. 3.2. In anaerobic condition no flux was measured through the enzyme, hence the reference condition for relative transcription differences was set to 21% oxygen level.

Flux through fumarase generally correlated with oxygen levels, its maximum was, however, reached at the second highest oxygen level (2.8%) rather than the highest. Quantitatively, transcript level data clearly correlated to the flux change. (Whether transcript levels at 2.8% and 21% were different is not conclusive from the data)

The following reversible Michaelis-Menten kinetic was used to calculate the flux through the enzyme

$$V^{\mathrm{FUM}} = \left(\frac{V_{\max}^{\mathrm{FUM}}}{Km_{\mathrm{FUM}}}\right) \frac{[FUM] \; - \; [MAL]/Keq}{1 \; + \; [FUM]/Km_{\mathrm{FUM}} \; + \; [MAL]/Km_{\mathrm{MAL}}} \tag{3.19}$$

$V_{\max}^{\mathrm{FUM}}$ fold change between reference condition (0% oxygen, denoted by subscript 0) and the other four conditions (denoted by subscript i) was estimated using the transcription amplification formula, Eq. 3.18:

$$\frac{V_i^{\max}}{V_0^{\max}} = \left(\frac{\mathrm{FUM1}_i}{\mathrm{FUM1}_0}\right)^{trexp} \tag{3.20}$$

hence the transcription amplification parameter *trexp* was incorporated as a parameter of the model to be fitted.

Regarding kinetic parameters, the only systematic study in an eukaryotic organism found (in the database BRENDA, [76]) used *Rattus norvegicus*. Hence we used this data to set search parameter ranges.

Fitting of calculated reaction rates to the measured flux values was performed by Monte-Carlo parameter search with 800 iteration repeated 20 times. Inspecion of the parameter stets from these repetitions yielded the following results:

parameter/quantity	search range	unit/remark	mean± stdev.
trexp	0 -7	exponent	3.8 ± 0.3
Km (fumarate)	0.013 - 0.333	mM	0.066 ± 0.038
Km (malate)	0.14 - 0.6	mM	0.56 ± 0.065
Keq	6 - 19	-	18.2 ±m 0.5
Vmax(0% O_2)	-	calculated	3.4 ± 0.9

Repetition of the estimation process showed that, even though some kinetic parameters exhibited considerable variance between estimations, the transcription amplification parameter *trexp* as well the quantities V, Vmax, g (c.f. Eq. 3.1), and therefore the flux dissection coefficients R_H, R_M, and R_{unk} (calculated via Eq. 3.8) proved remarkably stable. These results are presented in panel C - F of Fig. 3.2 for a parameter set found best in one execution of the estimation process. Analogous plots from any of the runs (not shown) do not differ significantly.

In accordance with the above mentioned correlation of transcription and flux, hierarchic coefficient R_H is the highest of the three coefficients with mean values between $\frac{1}{2}$ and 1. This expresses the fact that hierarchic regulation is assigned the largest contribution to the flux change under the model's assumptions, while R_M has values near or slightly below zero showing that metabolic regulation acts slightly antagonistic to the measured flux change, making it necessary for hierarchic regulation to compensate. This corresponds to the fact that substrate/product ratio change between conditions acts antagonistically to the flux change from the 0% condition; this ratio is lowest at 2.8% oxygen where measured flux is highest. This may be interpreted such that fumarase is hierarchically regulated under these conditions and actively contributes to the flux difference itself, thus influencing neighbouring reactions via metabolic regulation rather than passively regulated by other reactions.

In the cultures with 1% and 2.8% oxygen, R_H is lower and roughly equals R_{unk} with values around $\frac{1}{2}$, showing that only half of the flux change from the 0% condition can be accounted for by the model.

While the general trend in flux change is reproduced, the model cannot reproduce the zero flux measured in the anaerobic condition: the anaerobic rates resulting from the estimation process are around 1 mM/min, hence show less than a 10-fold change from the 21% O_2 condition.

This follows from the fact that, while, under the model's assumptions, hierarchic regulation (which we associate with transcription) is the only source for flux differences between conditions, the expression fold change between the 0% and 0.5% O_2 conditions is only slightly bigger than between other conditions.

This leads us to the interesting question whether the apparent lack of flux through fumarase at anaerobic condition is due to complete absence of active enzyme or due to equilibrium of substrate and product (which is not unlikely if this reaction essentially becomes a dead-end reaction under anaerobic conditions when its neighbour in TCA cycle, complex II, stops working). Since reproducing flux data required to set the equilibrium (Keq) relatively far from the measured substrate/product ratio, this provides an argument for the lack of enzyme activity being the main reason. An argument against this is the relatively high anaerobic transcript level with a fold difference of merely two, compared to the 21% condition.

A probable scenario consistent with the above is that the assumption of 'concerted regulation' (Eq. 3.16) is only an approximation, i.e. transcription change in this case differs somewhat from other simultaneous regulatory processes at various levels to regulate protein activity, which we can only partly reproduce from the information contained in *FUM1* transcription data. One possible source of the zero flux through this enzyme under anaerobic conditions is allosteric inhibition. Indeed, activation by acetate is reported in [64], c.f. Fig. 5.1. This offers an explanation to the above issue by assuming lower acetate levels in aerobic conditions and a corresponding effect on the mitochondrial enzyme under consideration.

From an evolutionary viewpoint, sustaining a certain level of enzymes associated to aerobiosis, even under prolonged anaerobic growth, may be interpreted as an investment in flexibility, since this strategy would enable the cell to activate respiratory metabolism faster, in case oxygen enters the system, possibly securing some advantage over competitors.

Phosphoglucose isomerase

Glucose-6-phosphate isomerase (EC 5.3.1.9) catalyses the second reaction of glycolysis. The only gene coding for this enzyme is PGI1 (CDC30, YBR196C).

It is generally assumed to be a 'fast' reaction with substrate-product ratio not far from equilibrium. It was shown that a 10-fold flux increase through the enzyme corresponded to only two-fold fold change in S/P-ratio during a double perturbation from glucose-limited aerobic to glucose-excess anaerobic condition [9].

As shown in 3.3, glycolytic flux through PGI decreased with oxygen provision, a generally known fact for budding yeast. The special case of a chemostat culture was discussed in Section 2.7. As the glycolytic flux between conditions differed

up to 5-fold, it is an interesting question, to what extent, based on the given data, this flux difference can be reproduced by our model assumptions.

Notably, substrate to product ratio did not differ significantly between conditions. This is in contrast to van den Brink *et al* [9], however not inconsistent with it, not only because glycolytic flux change was smaller in our experiments, but also because the measurements in [9] were conducted shortly after the perturbation, while in our case, they reflect the state of a continuous culture after many hours of steady-state growth.

Nevertheless, there were over two-fold differences in the absolute metabolite levels between conditions ('W-pattern', c.f. Fig. 2.3). This lead to similar pattern in the calculated reaction rates.

Transcription data exhibited only small changes and weak qualitative similarity to flux data: from anaerobic through the three microaerobic conditions (0.5%, 1% and 2.8%), *PGI1* transcription showed a weak decreasing tendency. However, at the 21% condition the mean value of the transcription data increases again to near-anaerobic level - even though high data variance leads to uncertainty about the exact level of relative upregulation in this condition.

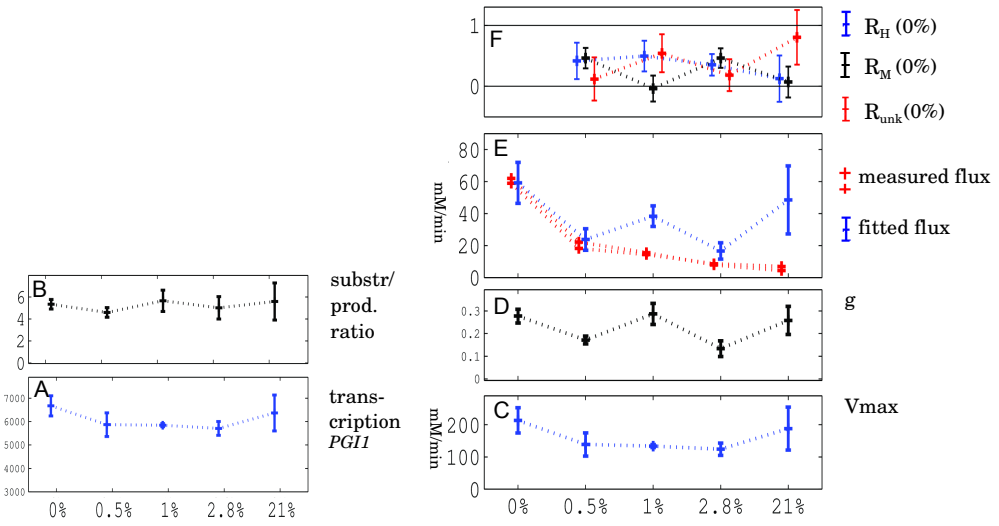

Figure 3.3: Measured quantities (panels A-B) and analysis (panels C-F) for PGI. Panel A shows and B show transcription of gene and substrate and product ratio. Panel C shows the activity V_{max} of the associated kinetics, as estimated from transcription data. Panel D shows the term g describing metabolite concentration effects in the kinetical expression. Panel E shows the calculated flux $V=V_{max}\, g(S)$. Panel F shows the flux difference analysis quantifying the contributions of metabolic, hierarchic, and unknown origins to flux change relative to the 0% experiment.
Calculations were based on single sample values, shown are means and standard deviations of the resulting statistical samples, except for experimental flux analysis for which both cultivations are shown separately.

Based on the Teusink model [87], reversible Michaelis-Menten kinetics was used to calculate the reaction rate:

$$V = \left(\frac{V_{\mathrm{PGI}}^{\max}}{Km_{\mathrm{G6P}}}\right) \frac{[G6P] \; - \; [F6P]/Keq}{1 \; + \; [G6P]/Km_{\mathrm{G6P}} \; + \; [F6P]/Km_{\mathrm{F6P}}} \tag{3.21}$$

Estimation of the change of $V_{\mathrm{PGI}}^{\max}$ between the reference condition (0% oxygen; subscript 0) and other conditions (subscript i) was again estimated using Eq. 3.18:

$$\frac{V_i^{\max}}{V_0^{\max}} = \left(\frac{\mathrm{PGI1}_i}{\mathrm{PGI1}_0}\right)^{trexp} \tag{3.22}$$

Kinetic parameters could be obtained in the database BRENDA [76], which allowed to set up parameter search ranges accordingly. A monte-carlo parameter search with 800 iteration performed 20 times yielded the following results:

parameter	search range	unit	result mean± stdev.
trexp	0 -5	(exponent)	3.8 ± 0.3
Km (G6P)	0.3 - 1.5	mM	$0.8 \pm 0.35(*)$
Km (F6P)	0 - 0.25	mM	0.23 ± 0.02
Keq	0.2 - 0.3	-	0.30 ± 0.0030
Vmax(0% O_2)	(calculated)	mM/min	311 ± 60

where $(*)$ denotes a parameter for which the estimation process did not result in significantly better limitations than that already defined by the search range (see Section 3.2.4). The rather tight boundaries for the equilibrium constant Keq are based on literature data. However, allowing larger variance did not improve the goodness of fit noticeably (not shown).

Figure 3.3 shows the measurement data and results of the fitting process and the flux dissection analysis. Calculation are analogous to those explained for fumarase. The measured flux data show 3-fold decrease between anaerobic and 0.5% oxygen conditions, and cultures with higher oxygen provision continue this tendency with a further three-fold difference between the 0.5% and the 21% cultures.

The achieved fit of reaction rates to the experimentally measured flux values is not satisfactory. The calculated reaction rates differ significantly from the measured flux values in the 1% and the 21% oxygen condition.

While a certain role of transcription in the flux difference between the anaerobic condition and the microaerobic conditions (0.5%, 1% 2.8%) seems probable from the qualitative similarity of the two data, the extent of it, as well as the role of metabolic regulation cannot be determined due to the inability of fitting measured fluxes,indicating inconsistency of data and the model assumptions.

In accordance to this, in spite of the relatively low variation in the S/P ratio, due to the great differences in absolute metabolite levels R_M varies greatly in these three conditions, the 1% culture showing much lower R_M value than the others.

Transcription does not play a role in the flux decrease from 2.8% to 21%) conditions, since *PGI1* is not downregulated further. Also the reported reactant level differences alone would not result in flux decrease. Hence, flux change

cannot be reproduced by the model (resulting in the high value of R_{unk} for this condition). Alone based on this, we would have to conclude that the observed further decrease in flux in this condition is either caused by post-transcriptional downregulation of PGI's activity (inconsistent with our model assumptions), or to the concentration change of one of the further substances which inhibit this enzyme (c.f. BRENDA entry for PGI), even though these are usually not regarded to assume a regulatory role for PGI.

However, it seemed to us, that failure of reproduction of the measured flux change may be traced back to the seemingly unintuitive pattern in reactant levels. A possible interpretation is to assume high, perhaps systematic measurement error in one or more of the cell cultures. Disregarding reported reactant levels open the possibility that the flux decrease in the 21% condition is due to metabolic regulation in connection with lower glucose levels - the latter is consistent with the conclusion in Section 2.7 on the residual glucose level being lowest in the aerobic condition.

We conclude that until new information becomes available, it remains principally unclear whether unsatisfactory reproduction of flux data by model reaction rates is due to principal limitations of the model (e.g. necessity of inclusion of further species) or due to experimental errors in the metabolite dataset.

A potential for measurement errors may be argued with the sensitivity of the system to quenching delay when harvesting samples from the culture, as discussed in Section 2.7. Concentration/flux ratio, a measure of how fast a certain reaction affects reactant levels, is $0.1 - 10$ seconds for glycolytic metabolites under these experiment's conditions. Since PGI is especially close to external glucose in the reaction chain, we expect that its reactant levels are rather susceptible to metabolic disturbances due to changes in the cells environment, especially the sudden drop in glucose levels when the sample is removed from the culture. Arguably, this may cause G6P and F6P level measurements especially sensitive to inhomogeneity of time delays during sampling.

For example, adopting the hypothesis that the quality of the 1% dataset is significantly lower than of other conditions, and omitting it from the analysis, one may arrive at the conclusion that flux changes from the reference condition(0%) to the microaerobic conditions (0.5%, 2.8%), are reasonably well reproduced, only change to aerobic condition is not. This is an intuitive conclusion, given the significant physiological differences between 2.8% and 21% oxygen cultures, but should be backed up with more data, for example new metabolite level measurements from analogous experiments.

Pyruvate dehydrogenase complex - gateway to the TCA cycle

As shown by the flux data in Fig. 2.2, there is a flux of pyruvate from glycolysis to TCA cycle even under anaerobic conditions when this pathway does not contribute to energy generation. In the mitochondria, oxidative decarboxylation of pyruvate occurs, catalysed by the large multienzyme complex known as pyruvate dehydrogenase complex (PDH-complex).

Pyruvate shuttling from cytosol to the mitochondria occurs by an active transporter the identity of which has long been unknown, and was only recently identified by Hildyard *et al.* [38] as a NAD transporter encoded by *YIA6*

(systematic name YIL006W) [86].

In eukaryotes, PDH-complex is localised in the mitochondrial matrix and consists of multiple copies of the following subunits: E1 or pyruvate dehydrogenase enzyme (EC 1.2.4.1, two subunits encoded by *PDA1, PDB1*), E2 or dihydrolipoyl transacetylase enzyme (EC 2.3.1.12, encoded by *LAT1*), E3 or dihydrolipoyl dehydrogenase enzyme (EC 1.8.1.4, encoded by *LPD1*), catalysing three sub-reactions respectively. A further component is Protein X or E3-binding protein (coded by *PDX1*) which plays a structural role in the complex by binding E3 to E2, the latter also playing a structural role as a central core of the complex [2], [76].

Being the gateway to aerobic metabolism, flux through the PDH-complex is subject to highly developed regulation. It is tightly regulated by its specific kinase/phosphatase pair: pyruvate dehydrogenase kinase (PDK, EC 2.7.11.2) and pyruvate dehydrogenase phosphatase (PDP, EC 3.1.3.43), which deactivate and activate the complex, respectively. Substrate and product of the PDH-complex affect these regulatory proteins which in turn effect the activity of the complex. The existence of this indirect regulatory mechanism seems necessary, since oxidative pyruvate decarboxylation is regarded as an irreversible reaction, hence its rate is not affected by product accumulation via purely kinetic mechanisms (metabolic regulation). The kinase/phosphatase pair is subject to its own hierarchic regulation: PDK is stimulated by ATP, NADH and acetyl-CoA and inhibited by ADP, NAD, CoA and pyruvate.

For this work, the above regulatory system was simplified into a single kinetics, since a more complex model would require more data. The inhibition by substrate via the kinase/phosphatase pair was simplified into an expression equivalent to allosteric inhibition. Since the metabolite quantification dataset did not include the product of the PDH-complex (Aceetyl-CoA), the next intermediate in the TCA-cycle, Citrate, was used in the role as inhibitor. The underlying assumption is that the ratio of Aceetyl-CoA and citrate does not vary greatly between conditions. Indeed, inclusion of citrate did somewhat improve the fit of the reaction rate to the experimentally determined flux.

Since NAD concentrations were not measured, we used the value used by Teusink, [NAD]=1.55 mM. The above considerations lead to the following 2-substrate irreversible Michaelis-Menten kinetics:

$$V = \frac{V^{\max}_{\mathrm{PDH}}}{(1 + \frac{[cit]}{Kphen_{\mathrm{cit}}})} \cdot \frac{\frac{[PYR]}{Km_{\mathrm{PYR}}} \frac{[NAD]}{Km_{\mathrm{NAD}}}}{(1 + \frac{[PYR]}{Km_{\mathrm{PYR}}})(1 + \frac{[NAD]}{Km_{\mathrm{NAD}}})} \tag{3.23}$$

where the parameter $Kphen_{\mathrm{cit}}$ denotes a phenomenological inhibition constant.

Fold change of $V^{\max}_{\mathrm{PDH}}$ between the reference condition (0% oxygen; subscript 0) and other conditions (subscript i) was estimated from transcription data, using the following expression:

$$\frac{V^{\max}_{\mathrm{PDH}\,i}}{V^{\max}_{\mathrm{PDH}\,0}} = \frac{YIA6_i^{trexp} + (PDA1 + PDB1 + LPD1 + LAT1)_i^{trexp}}{YIA6_0^{trexp} + (PDA1 + PDB1 + LPD1 + LAT1)_0^{trexp}} \tag{3.24}$$

Results from a monte-carlo parameter search with 800 iterations are shown in Fig. 3.4. A statistical evaluation of the search process performed 20 times

yielded the following results:

parameter	search range	unit	result mean± stdev.
trexp	0 - 5	(exponent)	4.8 ± 0.12
Km (pyruvate)	0 - 0.01	mM	0.002 ± 0.0013
Km (NAD)	0.5 - 3	mM	$2.0 \pm 0.7\,(*)$
Kphen (citrate)	2 - 6	-	5.9 ± 0.14
Vmax(0% O_2)	(calculated)	mM/min	15.4 ± 3

where $(*)$ denotes a parameter for which the estimation process did not result in significantly better limitations than that defined by the search range (see Section 3.2.4).

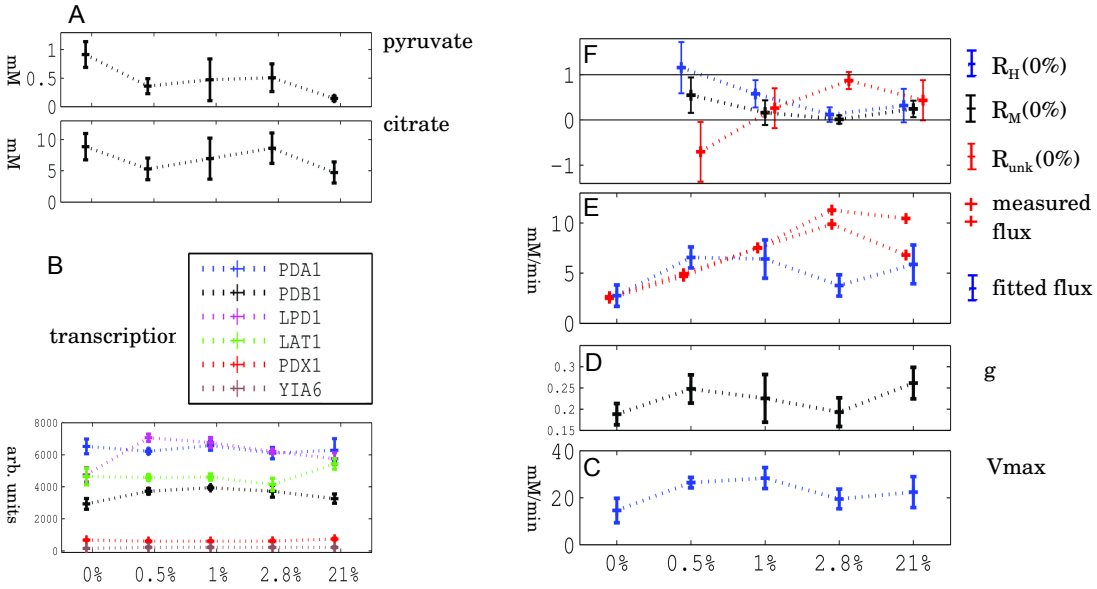

Figure 3.4: Measured quantities (panels A - B) and analysis (panels C-F) for the PDH-complex. *Panel A* shows concentrations of pyruvate, the substrate, and citrate which was taken as an effector. *Panel B* shows transcription of the genes used in the calculation of V_{max}. *Panel C* shows the activity V_{max} of the associated kinetics, as estimated from transcription data. *Panel D* shows the term g describing metabolite concentration effects in the rate expression. *Panel E* shows the calculated flux $V=V_{max}\, g(S)$. *Panel F* shows the flux difference analysis quantifying the contributions of metabolic, hierarchic, and unknown origins to flux change relative to the 0% experiment.
Calculations and fitting were based on single sample values.

Results of one parameter search process is shown in Fig. 3.4. While the basic qualitative tendency of the flux data - flux increase with higher oxygen provision - is reproduced by the calculated reaction rate, finer details are only reproduced to a low degree: while the experimental data show a clear flux increase with higher O_2 provision up to 2.8%, then a slightly lower value for the 21% condition, the only flux difference qualitatively reproduced by the kinetics

is the rise from the anaerobic to 0.5% condition. Nevertheless, absolute flux values are approximately reproduced in four of the five conditions, however, this stems from the fact that the notable reaction rate difference produced by the model between anaerobic and 0.5% conditions is 1.5 times larger than would be necessary for the measured flux data (indicated by the corresponding R_{unk} value of -0.5).

Inspection of pyruvate levels reveals that flux change can not be induced by substrate accumulation, since pyruvate levels show, in contrast to flux data, a decrease with higher oxygen levels. Still, the (larger than necessary) change from anaerobic to 0.5% is associated with a relatively high R_M mean value of ca. 0.5, half of the R_H value, indicating that metabolic regulation still caused around one third of the reproduced rate difference. Since NAD level is a constant in this model, this stems from the fact that the relative drop of pyruvate is apparently more than counterbalanced by the drop in citrate level (playing the role of an inhibitor in this model).

The increase of measured flux from 1% to 2.8% condition is not reproduced by the model. Visual inspection of the experimental data reveals that they contain hardly any information to explain this flux increase. Clearly, for this to be explained, new data input is necessary.

Flux dissection analysis shows for the flux change from anaerobic to 0.5%, and to 1% conditions relatively high mean R_H values (c.a. 1 and 0.6 respectively), indicating that the calculated changes in the reaction rate stem to a high extend from hierarchic regulation. While R_H values for the other two aerobic condition are higher than R_M values, this is diminished by high R_{unk} values, ie. the fact that these flux changes are not well reproduced by the model.

Overall, the data indicate a significant role of both metabolic and hierarchic regulation in the flux difference between the states in anaerobic and aerobic conditions, however no strong evidence for the source of fluxes between the various aerobic conditions is found. Due to the low quantitative reproduction of the flux change by the calculated reaction rate, flux dissection analysis is found of limited use in this case.

The question arises whether this lack of reproducing more than the most basic pattern of flux change is due to oversimplification, possibly trying to fit the elaborate regulatory mechanism of PDH-complex into a relatively simple kinetics. However, inspection of the input data (panels A and B in Fig. 3.4) lead us to the conclusion that it is by no means obvious which combination of these data would reproduce the further increase in reaction rate seen in the measured flux data, even if a more complex model were used to interpret it.

Of course, inherent to our approach is that post-transcriptional mechanisms, not captured by the available data, and not in concerted action with transcription (see Eq. 3.16), impair the ability of the model to reproduce patterns in flux change.

Aside from this, we can identify two potential reasons. Firstly, the falsity of the assumption that citrate has a similar steady state concentration pattern to Acetyl-CoA, the immediate product and important effector of the complex. On the other hand, this may be an indication for a substantial metabolic control of the flux through that reaction

Secondly, it may well be that more data, most notably information on the

state of the kinase-phosphatase mediated regulatory mechanism is necessary for reproducing flux patterns under these experimental conditions. Ideally this would include information on the number of (functional) PDH-complexes, the ratio of active to inactivated (phosphorylated) complexes, ratio of PDP and PDK regulatory enzymes, as well as information on their effectors.

3.3.2 PFK - FBP: the switching circuit of upper glycolysis

The reaction catalysed by the enzyme phosphofructokinase is usually regarded as on of the important control points in the metabolic flux. Since F6P, the substrate of PFK is connected to G6P by the reversible reaction (catalysed by PGI), and G6P is a branching point for a number of pathways, most notably the pentose-phosphate-pathway, changing the net rate of the PFK-FBP system is an effective way for metabolic regulation of the ratio between the main glycolytic flux, and through the reactions branching from upper glycolysis,

Since it influences the concentration of G6P, it is in the position to regulate flux distribution between the main glycolytic flux and all processes branching off of upper glycolysis, such as the pentose phosphate pathway. (As discussed above, the enzyme PGI is close to in equilibrium, thus readily mediating F6P accumulation backwards.)

The enzyme PFK itself is regulated by numerous signals in a complex way - it has been called "the modellers nightmare" by Teusink [87]. However, in order to understand the allosteric regulatory network, it is not sufficient to concentrate only on the enzyme PFK, rather it should be viewed in the context of the regulatory unit shown in Fig. 3.5, consisting of the two pairs of antagonistic enzymes phosphofructokinase (PFK) - fructose-bisphosphatase (FBP), and phosphofructokinase 2 (PFK2) - fructose-2,6-bisphosphatase (FBP2), respectively.

The reaction catalysed by PFK consumes ATP, hence, this reaction is well-known to be irreversible under physiological conditions. The reverse reaction is catalysed by fructobisphosphatase (FBP, encoded by gene FBP1), which does not involve cofactor, and is known to facilitate gluconeogenesis. The enzyme-pair is known as a 'potential futile cycle' and, to our knowledge, there is no general consensus on whether or under what condition both enzymes are active at the same time causing some loss of free energy (ATP) into heat.

The metabolite F26bP plays an important role in this context, since it mediates the effect of Phosphoenolpyruvate (PEP), hence carries information on the state of lower glycolysis. Indeed, differences between PEP levels were significant between conditions.

In addition to the allosteric regulatory interactions shown in Fig. 3.5, the involved enzymes are influenced by the concentration of a number of metallic ions as well as by phosphorylation by cAMP-dependent protein kinase (PKA), hence also influenced by cAMP level. Given the limited amount of data, typically only a part of known regulatory interactions can be represented in a model, and certain species are assumed to be constant.

Naturally, not all species playing a role in the above outlined system were represented in the dataset, shown in Eq. 3.6
On the other hand, even if data is available, the complexity of the model (and number of parameters) would rise with each included species.

We aimed to include information on the state of lower glycolysis in the kinetics rate expression. In the above outlined regulatory system, this information is mediated on the metabolic level via PEP and the TCA intermediate, citrate. The allosteric effect of F26bP on PFK was already inlcuded in the detailed model of Teusink. Since this is the metabolite carrying the information of PEP levels into the regulation of the PFK module, we decided to link PEP levels to F26bP in the model. In order to estimate changes in F26bP level, we extended the system by including a simpler steady-state model of the reactions catalysed by the regulatory enzymes PFK2 and FBP2, including the effect of PEP and their respective transcription.

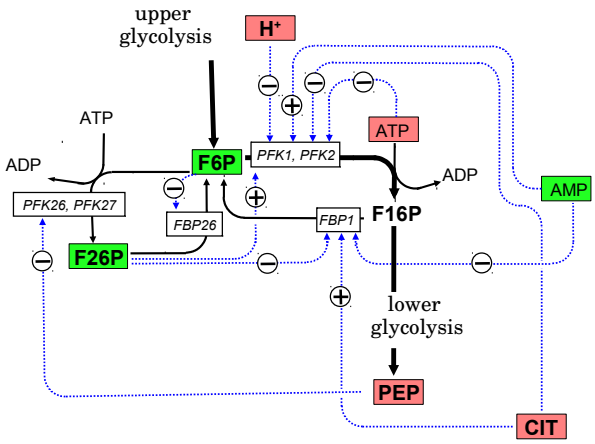

Figure 3.5: Allosteric regulatory interactions of the PFK-FBP unit.
Species whose accumulation, causes net glycolytic flux through PFK-FBP to decline, are marked red; those whose accumulation leads to an increase, are marked green. Thick arrows mark the main glycolytic flux in glucose consuming mode.
PFK (encoded by genes *PFK1* and *PFK2*) and FBP (*FBP1*) may work antagonistically in yeast under aerobic sugar growth conditions. The futile cycle composed by PFK2 (*PFK26, PFK27*) and FBP2 (*FBP26*) has a regulatory role via regulating the concentration of F26bP.
According to the diagram, accumulation of upstream species or AMP (associated with low ATP levels) results in increase of the net flux, while accumulation of the downstream species phosphoenolpyruvate (PEP) and Citrate (CIT) or ATP result in a decrease of the flux.
The model for the PFK unit includes all four enzymes indicated and omits the allostiric modifiers Citrate, PH (H+) and F6P. (Diagram modified from [64])

A model for the PFK-FBP regulatory unit

The model used for this model is based on the PFK-kinetics constructed by Teusink [87], one of the most complex single kinetics expression in the literature used for a single enzyme. This kinetics was complemented by a kinetical expression for the gluconeogenetic enzyme FBP and an expression to estimate the F26bP steady-state levels based on kinetical expressions for the enzymes PFK2 and FBP2.

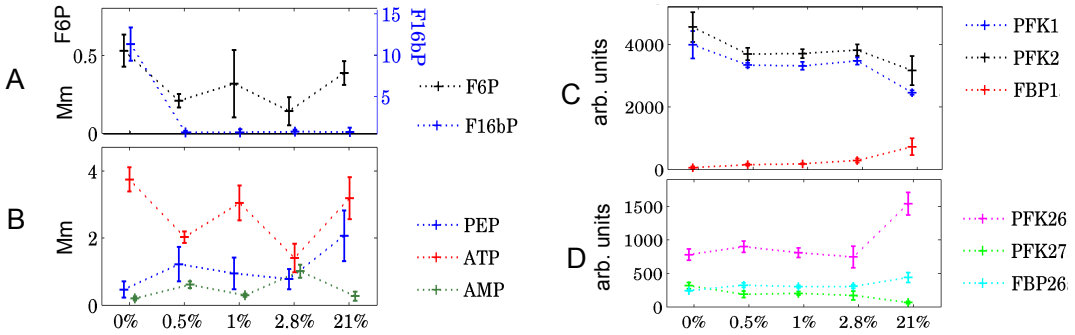

Figure 3.6: Data input used for modelling the PFK - FBP module (c.f. Fig. 3.5). Panel A and C show metabolite and expression data corresponding to the enzymes associated with the main branch of glycolysis, PFK (*PFK1*, *PFK2*), and FBP (*FBP1*).
Panel B and D show metabolite and expression data corresponding to the regulatory branch consisting of the enzymes PFK2 (*PFK26*, *PFK27*), and FBP2 (*FBP26*).

Due to the complexity of this system it was not attempted to undertake a parameter fit. Since the complexity of the model was kept to a minimum, all parameter values could be taken from the Teusink model, found in the literature as *in-vitro* data, or calculated from other data as described below.

Naturally, a number of regulatory interactions were not included in the model, notably the pH-sensitivity of PFK (creating a regulatory connection to the pentose-phosphate-pathway and other reactions changing the pH), and the Citrate sensitivity of PFK and FBP (carrying information on the state of the TCA-cycle) were omitted, partly due to lack of data, but also in order not to overload an already complicated model.

The enzyme PFK is modelled in the Teusink-model by the following kinetics expression (based on the Monod - Wymann - Changeux model for allosteric enzymes, as adapted by Hess and Plesser, s. [87]) This kinetics model accounts for the allosteric inhibitory effect of F16bP and ATP as well as the activation by F26bP and AMP.

$$V^{\mathrm{PFK}}(F6P, F16bP, F26bP, ATP, AMP) = \\ = V^{\mathrm{PFK}}_{max}\, g_R\, \frac{F6P}{K_{R,\mathrm{F6P}}}\, \frac{ATP}{K_{R,\mathrm{ATP}}}\, \frac{R}{R^2 + LT^2} \tag{3.25}$$

$$\begin{aligned} \text{where } R &:= 1 + \frac{F6P}{K_{R,\mathrm{F6P}}} + \frac{ATP}{K_{R,\mathrm{ATP}}} + g_R \frac{F6P}{K_{R,\mathrm{F6P}}} \frac{ATP}{K_{R,\mathrm{ATP}}} \\ T &:= 1 + C_{i\mathrm{ATP}} \frac{ATP}{K_{R,\mathrm{ATP}}} \\ L &:= L_0 \left(\frac{1 + Ci_{\mathrm{ATP}} \frac{ATP}{K_{\mathrm{ATP}}}}{1 + \frac{ATP}{K_{\mathrm{ATP}}}} \right)^2 \cdot \left(\frac{1 + C_{i\mathrm{AMP}} \frac{AMP}{K_{\mathrm{AMP}}}}{1 + \frac{AMP}{K_{\mathrm{AMP}}}} \right)^2 \cdot \\ & \qquad \left(\frac{1 + C_{i\mathrm{F26bP}} \frac{F26bP}{K_{\mathrm{F26bP}}} + C_{\mathrm{F16bP}} \frac{F16bP}{K_{\mathrm{F16bP}}}}{1 + \frac{F26bP}{K_{\mathrm{F26bP}}} + \frac{F16bP}{K_{\mathrm{F16bP}}}} \right)^2 \end{aligned}$$

abbr.	name, E.C.	gene(s)	equation parameter	value	unit	reference
PFK	Posphofructokinase, 2.7.1.11	*PFK1, PFK2*	Eq. 3.25			[87]
			$V_{\max}^{\mathrm{PFK}}$	s. Table 3.3	mM/min	
			$K_{R,\mathrm{F6P}}$	s. Table 3.3	mM	[87]
			$K_{R,\mathrm{ATP}}$	0.71	mM	[87]
			g_R	5.12	-	[87]
			L_0	0.66	-	[87]
			C_{ATP}	3	-	[87]
			Ci_{ATP}	100	-	[87]
			K_{ATP}	0.65	mM	[87]
			Ci_{AMP}	0.0845	-	[87]
			K_{AMP}	0.0995	mM	[87]
			Ci_{F16bP}	0.397	-	[87]
			K_{F16bP}	s. Table 3.3	mM	[87]
			Ci_{F26bP}	0.0174	-	[87]
			K_{F26bP}	s. Table 3.3	mM	[87]
FBP	fructose-bisphosphatase, 3.1.3.11	*FBP1*	Eq. 3.26			
			$V_{\max}^{\mathrm{FBP}}$	s. Table 3.3	mM/min	
			Ki_{F26bP}	s. Table 3.3	mM	
			Km_{F16P}	s. Table 3.3	mM	
PFK2	6-phosphofructo-2-kinase, 2.7.1.105	*PFK26 PFK27*	Eq. 3.30			
			V_{max}^{PFK2}	not used	mM/min	
			Km_{F6P}	0.4	mM	BR:[3]
			Ki_{PEP}	2	mM	BR:[28]
FBP2	fructose-2,6-bisphosphate 2-phosphatase, 3.1.3.46	*FBP26*	Eq. 3.30			
			V_{max}^{PFK2}	not used	mM/min	
			Km_{F26bP}	0.002	mM	BR:[28],[52]

Table 3.2: List of parameters and represented enzymes in the model of the PFK - FBP regulatory unit. For maximal velocities of the enzymes PFK2 and FBP2, only the ratio ($\frac{V_{max}^{\mathrm{PFK2}}}{V_{max}^{\mathrm{PFK2}}}$) was used, c.f. Table 3.3 and Fig. 3.33. Citations preceded by BR were found in the database BRENDA [76].

Most parameters were taken from [87] (s. Table 3.2), except for V_{max}^{PFK} which was adjusted as explained below.

The gluconeogenetic enzyme FBP was modelled with an irreversible Michaelis-Menten kinetics assuming non-competitive inhibition by F26bP.

$$V^{\mathrm{FBP}}(F16bP, F26bP) = \left(\frac{1}{1+\frac{F26bP}{Ki_{\mathrm{F26bP}}}}\right) \frac{V_{max}^{\mathrm{FBP}}}{1+\frac{Km_{\mathrm{F16P}}}{F16P}} \tag{3.26}$$

The enzyme activities V_{max}^{FBP} and V_{max}^{PFK} were - in a slight difference to the algorithm explained before - estimated as follows (Experimental conditions are denoted by subscripts showing the oxygen percentage):

The value of $V_{max\,21\%}^{\mathrm{FBP}}$ was represented as a free parameter. Fold changes from 21% oxygen to the other four conditions (denoted by subscript i) were

calculated from transcription data of the enzyme $FBP1$, using the transcription amplification parameter *trexp* (c.f. Eq. 3.18):

$$\left(\frac{V^{\mathrm{FBP}}_{max\,i}}{V^{\mathrm{FBP}}_{max\,21\%}}\right) = \left(\frac{FBP1_i}{FBP1_{21\%}}\right)^{trexp} \tag{3.27}$$

Hence, for each iteration of the parameter search with a new set of parameters, the two parameters *trexp* and $V^{\mathrm{FBP}}_{max\,21\%}$ determine the enzyme activity of FBP for each experimental conditions. The 21% oxygen provision condition was chosen as reference state for this case for practical reasons: FBP exhibits a large upregulation with higher oxygen provisions, with near-zero anaerobic transcription level, presumably corresponding to practically zero enzymatic activity. Hence the biologically relevant value in the 21% condition was chosen as reference.

For the enzyme PFK, we could use the anaerobic condition as reference. For given parameter values, and $V^{\mathrm{FBP}}_{max\,21\%}$ calculated from Eq. 3.27, the value of $V^{\mathrm{PFK}}_{max\,21\%}$ was determined by the condition that the net calculated reaction rate matched the measured flux through this part of glycolysis:

$$\mathrm{mean}(V^{\mathrm{PFK}}_{max\,0\%} - V^{\mathrm{FBP}}_{max\,0\%}) = \mathrm{mean}(J_{exp\,0\%}) \tag{3.28}$$

where $J_{exp\,0\%}$ denotes the measured flux sample values for the anaerobic condition, and *mean* is taken over the single sample values of these quantities (c.f. Section 3.2.3). We note that in the scenarios presented in this chapter, the value $V^{\mathrm{FBP}}_{max\,0\%}$ was negligible in comparison to PFK's activity.

Subsequently, the fold change of V^{PFK}_{max} between conditions was determined from the average transcription fold change of the genes encoding the two PFK subunits:

$$\left(\frac{V^{\mathrm{PFK}}_{max\,i}}{V^{\mathrm{PFK}}_{max\,21\%}}\right) = mean\left(\frac{PFK1_i}{PFK1_{0\%}}, \frac{PFK2_i}{PFK2_{0\%}}\right)^{trexp} \tag{3.29}$$

Note that the same value for the transcription amplification parameter *trexp* was used in Eq. 3.27 and 3.29. Hence, in the above described scheme, two parameters (*trexp* and $V^{\mathrm{FBP}}_{max\,21\%}$) determine the values for the maximal velocities of PFK and FBP, except for $V^{\mathrm{PFK}}_{max\,0\%}$ which is determined by Eq. 3.28.

A steady-state model the PFK2-FBP2 branch for predicting F26bP concentrations

As mentioned above, levels of the regulatory metabolite F26bP were not measured, hence we attempted to estimate its levels by a comparatively simple two-enzyme-model using corresponding transcription data and available literature values. This enzyme-kinetics model, described in this section, consists of two irreversible Michaelis-Menten reactions in which the inhibitory effect of PEP on PFK2 is modelled by a non-competitive inhibition term:

$$V^{\mathrm{PFK2}}(F6P, PEP) = \left(\frac{1}{1+\frac{PEP}{Ki_{\mathrm{PEP}}}}\right) \frac{V_{max}^{\mathrm{PFK2}}}{1+\frac{Km_{\mathrm{F6P}}}{F6P}} \tag{3.30}$$

$$V^{\mathrm{FBP2}}(F26bP) = \frac{V_{max}^{\mathrm{FBP2}}}{1+\frac{Km_{\mathrm{F26bP}}}{F26bP}} \tag{3.31}$$

The relative simplicity of these equations and the regulatory term is again in correspondence with the amount of information available for this specific system. Our conjecture that a two-reaction model would estimate the changes in F26bP level is based on the assumption that PFK2 and FBP2 are the only two specific enzymes in yeast with F26bP as a reactant. That is consistent[5] with the databases KEGG [49] [48], and SGD [1]. This assumption is insofar an approximation, that a number of other enzymes may (non-specifically) hydrolyse this metabolite, however their affinity was reported to be too low ($Km > 0.05$ mM) for the concentrations in this experiment. [28].

Parameters for the kinetic parameters could be found in the literature.

The assumption of steady state

$$V^{\mathrm{PFK2}}(F6P, PEP) = V^{\mathrm{FBP2}}(F26bP) \tag{3.32}$$

allows - due to the simplicity of the involved kinetical expressions - to express the steady-state concentration $F26bP^*$

$$F26bP^* = \frac{Km_{\mathrm{F26bP}}}{\left(\frac{V_{max}^{\mathrm{FBP2}}}{V_{max}^{\mathrm{PFK2}}}\right)(1+\frac{PEP}{Ki_{\mathrm{PEP}}})(1+\frac{Km_{\mathrm{F6P}}}{F6P}) - 1} \tag{3.33}$$

The term $\left(\frac{V_{max}^{\mathrm{FBP2}}}{V_{max}^{\mathrm{PFK2}}}\right)$ is still unknown at this point. In order to estimate its values for the different conditions, we relate it to a reference condition with known concentrations $\mathrm{F6P}_0$, $\mathrm{F26bP}_0$, and PEP_0 in which case we can use again Eq. 3.32 to gain the expression to calculate it as

$$\left(\frac{V_{max}^{\mathrm{FBP2}}}{V_{max}^{\mathrm{PFK2}}}\right)_0 = \frac{(1+\frac{Km_{\mathrm{F26bP}}}{F26bP_0})}{(1+\frac{PEP_0}{Ki_{\mathrm{PEP}}})(1+\frac{Km_{\mathrm{F6P}}}{F6P_0})} \tag{3.34}$$

This allows to estimate $\left(\frac{V_{max}^{\mathrm{FBP2}}}{V_{max}^{\mathrm{PFK2}}}\right)$ for a reference condition, using the presented PEP and F6P concentrations in combination with a matching literature value for F26bP concentration.

In accordance with our general approach, we chose to select the anaerobic condition as reference state, using the concentration $F26bP_0 = 0.020$ mM/l reported in [87].

In the second step we estimate the fold change to $\left(\frac{V_{max}^{\mathrm{FBP2}}}{V_{max}^{\mathrm{PFK2}}}\right)$ for the other four conditions, using transcription data for the genes coding for these enzymes.

[5] According to the KEGG database, the enzyme fructose-2,6-bisphosphate-6-phosphatase, EC 3.1.3.54 is not present in budding yeast. The gene YLR345Wp has high similarity to these enzymes but at this point there is no experimental evidence for its role [100].

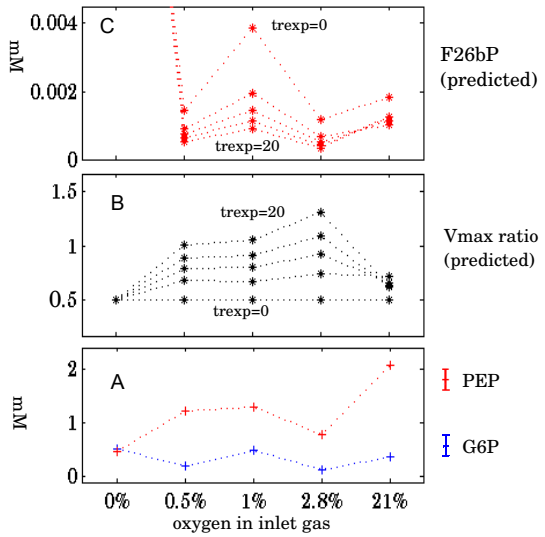

Figure 3.7: F26bP concentration predicted for aerobic conditions by the steady state model using metabolic concentration and transcription data.
Panel A shows experimental data for F6P and PEP concentrations. In addition, a literature value for F26bP was used for the 0% oxygen reference condition.
Panel B shows the predicted enzyme activity fold change $\left(\frac{V_{max}^{\mathrm{FBP2}}}{V_{max}^{\mathrm{PFK2}}}\right)$ as calculated from metabolite (reference condition) and transcription data (aerobic conditions) using Eq. 3.35 with the following values of the parameter *trexp*: 0 (no transcriptional regulation), 5, 10, 15, and 20.
Panel C shows the corresponding predicted steady state F26bP concentrations, calculated using Eq. 3.33. The value for the reference condition is not shown since, by construction, it is the literature value [F26bP]= 0.02 mm/L for all values of *trexp*.
In spite of the slight variations of the predicted aerobic values with *trexp*, all parameter values predict the drop of concentration from anaerobic to aerobic conditions by an order of magnitude.

$$\begin{aligned}
\left(\frac{V_{max}^{\mathrm{FBP2}}}{V_{max}^{\mathrm{PFK2}}}\right)_i &= \left(\frac{V_{max}^{\mathrm{FBP2}}}{V_{max}^{\mathrm{PFK2}}}\right)_0 \cdot \frac{F_{tr\,i}^{\mathrm{FBP2}}}{F_{tr\,i}^{\mathrm{PFK2}}} \qquad (3.35)\\
\text{with } F_{tr\,i}^{\mathrm{PFK2}} &= \frac{PFK26^{trexp} + PFK27^{trexp}}{PFK26_0^{trexp} + PFK27_0^{trexp}}\\
F_{tr\,i}^{\mathrm{FBP2}} &= \frac{FBP26^{trexp}}{FBP26_0^{trexp}}
\end{aligned}$$

where $F_{tr\,i}^{\mathrm{PFK2}}$ and $F_{tr\,i}^{\mathrm{FBP2}}$ denote the estimated fold changes for the respective enzyme activity, calculated from the transcription data, denoted by the respective gene's name; the exponent *trexp* is a free parameter (c.f. Eq. 3.18).

Note that this approach does not require to obtain absolute values for these enzyme activities.

Figure 3.7 shows the resulting prediction for F26bP concentration for the four aerobic states. It was found that, under the assumptions of the above model, this does not depend crucially on the extent of enzyme activity fold change, hence the value for the parameter *trexp* is of little importance for the range under study. For all shown values, the estimated F26bP concentration for the 21 % oxygen condition is of the same order of magnitude as the concentration reported in [9]. From the reported value 0.002 μmol/gDW we estimated a concentration of 0.0013 mM. Experimental conditions in this study (steady state culture in glucose limited chemostat with D=0.1/h) as well as yeast strain (CEN.PK113-7D, MATa) closely resembled our aerobic experiments.

This behaviour essentially stems from the reported value of the FBP2 parameter Ki_{PEP} which makes the system especially sensitive to PEP concentration change for the range which occured in the experiment. On the other hand, the used PFK kinetics does show sensitivity to F26bP changes in this range (0.0013-0.02 mM) under certain conditions. (This property depends on the parameters Ci_{F26bP} and K_{F26bP}.)

In this sense, the parameter values from different sources fit together to produce sensitivity of the PFK kinetics for the reported range of PEP levels.

Results and discussion

In the above described model for the PFK-unit, data input shown in Fig. 3.6 was used. Monte-Carlo parameter search was applied to study how well the model was able to fit the net flux $\mathrm{V}^{\mathrm{net}} = \mathrm{V}^{\mathrm{PFK}} - \mathrm{V}^{\mathrm{FBP}}$ to the measured values. Flux dissection analysis was carried out for the net flux of the enzymes PFK and FBP, making use of the linear nature of the coefficients defined in Eq. 3.8 allowing to simply add R_M, R_H, and R_{unk} for the two enzymes. Results from three parameter search algorithms with 3000 iterations is shown in Fig. 3.8. Results using 20 repeats of parameter search are presented in Table 3.3.

Generally, transcription activity patterns between conditions (Fig. 3.6) showed high consistence with the biochemical knowledge represented in the regulatory network in Fig. 3.5 and the measured lower glycolytic flux at higher oxygen provision levels. The genes *PFK1-2*, and *PFK26-27*, encoding enzymes which exert a direct or indirect positive effect on the net glycolytic flux (c.f. Fig. 3.5), generally showed lower transcription in conditions with higher oxygen levels, while the enzymes *FBP1* and *FBP26* showed upregulation.

On the metabolite level, the lower-glycolytic effector PEP (acting as an indirect inhibitor of the net glycolytic flux through this reaction) further contributed to this trend, exhibiting higher levels in conditions with higher oxygen. Combined with the effect of expression patterns of the genes encoding PFK2 and FBP2, the predicted levels of the effector metabolite F26bP (s. Fig. 3.7) also indicate downregulation of the net glycolytic flux.

Interestingly, most transcription patterns seem to exhibit three distinct phases: significant differences mainly exist between anaerobic, microaerobic and fully aerobic conditions, but not between the three microaerobic conditions (0.5-2.8% oxygen). *FBP1* expression is a notable exception to this as significant transcription levels are only shown for the two highest oxygen provision conditions. On metabolite level, PEP exhibits this trend.

parameters	range	$\text{trexp}_{\max} = 5$	$\text{trexp}_{\max} = 2$	$\text{trexp}_{\max} = 1$
trexp	0 - $\text{trexp}_{\max}$	4.6 ± 0.20	1.9 ± 0.06	0.94 ± 0.04
$V_{\max}^{\text{FBP}}(21\%)$	0 - 100	1.1 ± 1	10 ± 2.4	26 ± 7.5
$\text{Km}_{\text{F16bP}}(\text{FBP})$	0.001 - 0.01	$0.006 \pm 0.003(*)$	$0.006 \pm 0.0024(*)$	$0.006 \pm 0.003(*)$
$\text{Ki}_{\text{F26bP}}(\text{FBP})$	0.001 - 0.03	$0.013 \pm 0.0089(*)$	$0.015 \pm 0.008(*)$	$0.01 \pm 0.008(*)$
$K_{R,\text{F6P}}(\text{PFK})$	0.03 - 0.33	$0.18 \pm 0.09(*)$	$0.19 \pm 0.1(*)$	$0.17 \pm 0.09(*)$
$\text{K}_{\text{F16bP}}(\text{PFK})$	0.03 - 0.33	0.29 ± 0.03	0.32 ± 0.01	0.32 ± 0.01
$\text{K}_{\text{F26bP}}(\text{PFK})$	0.0001 - 0.001	$0.0006 \pm 0.0003(*)$	$0.0005 \pm 0.0002(*)$	$0.0005 \pm 0.0003(*)$
calculated quantities				
$V_{\max}^{\text{PFK}}(0\%)$	-	104.5 ± 3.2	103 ± 0.9	104 ± 0.8
$V_{\max}^{\text{PFK}}(21\%)$	-	15.0 ± 1.5	45.6 ± 1.1	69.8 ± 1.3

Table 3.3: Results of the parameter estimation procedure of the PFK regulatory unit with three different values for $trexp_{\max}$, the upper limit of the transcription amplification parameter *trexp*. (c.f. Fig. 3.8) For each value of $trexp_{\max}$ a monte-carlo parameter search with 3000 iterations was performed 20 times. Mean and standard deviation (std) are shown (rounded values). Parameters for which standard deviation is in the order of one third of the search range, are assumed to not further determined by the estimation process, and are marked with $(*)$.
For the kinetic parameters K_{F6P} K_{F16bP} K_{F26bP} for PFK, the search ranges were defined by [3.3 c_T , $c_T/3.3$] with c_T denoting the value used by Teusink. Concentration unit is mM, reaction rate unit is mM/min.

Net glycolytic fluxes through the reaction catalysed by PFK-FBP were reproduced mostly satisfactorily in terms of absolute values. The general trend, continuous decrease of net flux towards higher oxygen levels, was, however, not well reproduced by the best-fit parameterset when the search limit for the transcription amplification parameter *trexp* was set to 5 (used as standard value).

The model exhibits notable flexibility with regard to the extent of hierarchic regulation in the sense that reduced changes in transcription may be counterbalanced by metabolic regulation in reproducing flux changes. To demonstrate and explore this flexibility, the parameter search was repeated with three different upper search limits for the transcription amplification parameter *trexp*: Fig. 3.8 shows results with $trexp_{max} = 5$, 2 , and 1.

Remarkably, using lower $trexp_{max}$ values resulted not only in good reproduction of the net flux values by calculated mean reaction rate values, but, what is more interesting, also in a better reproduction of the above-mentioned trend of flux decrease between conditions. Variance of single sample values (shown as error-bars in Fig. 3.8) increased in these fits, leading to higher objective function values (c.f. Eq. 3.19).

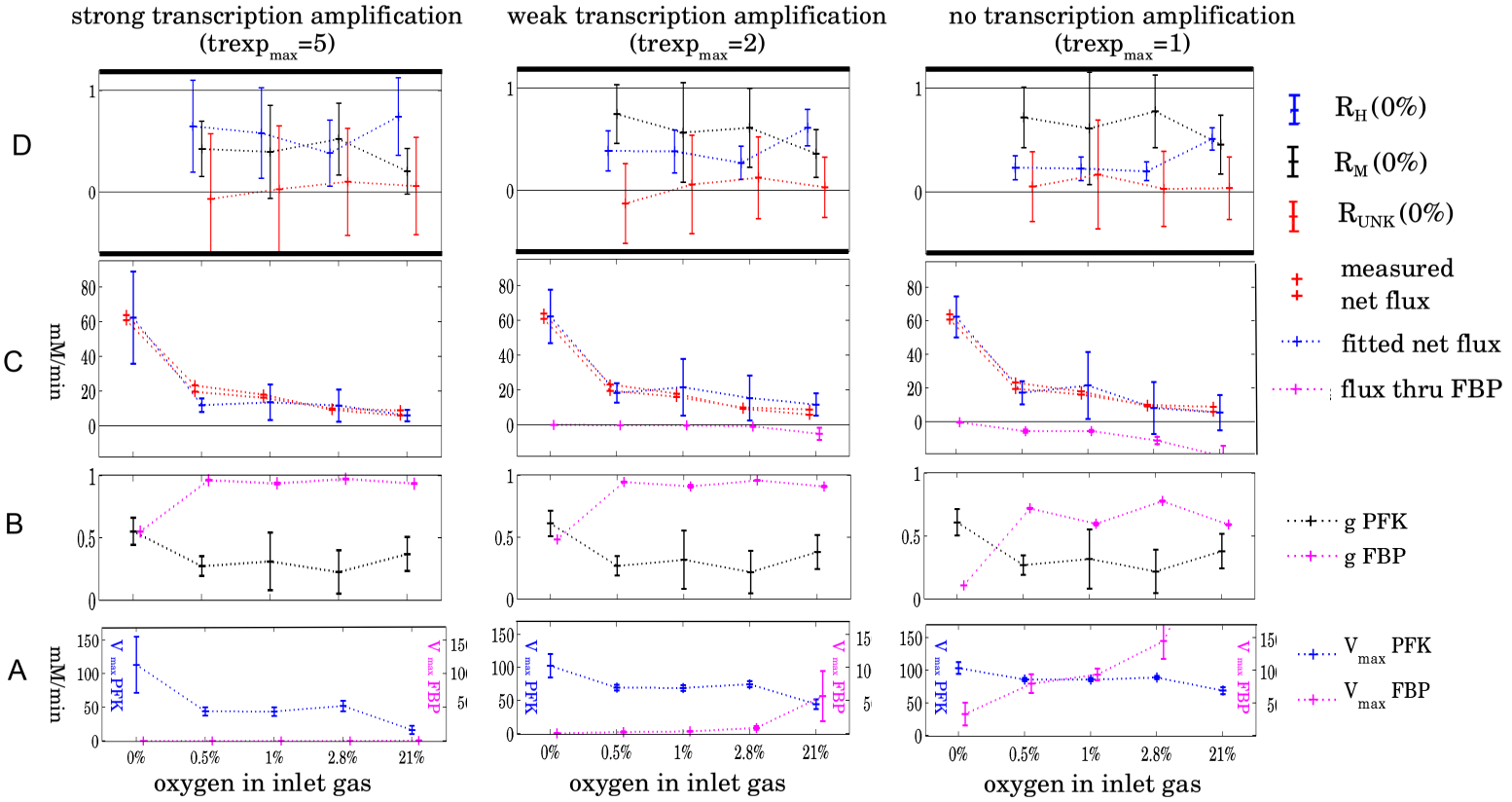

Figure 3.8: Results of fitting the model reaction rate of the PFK - FBP unit to the measured net flux varying the strength of transcription amplification. Plots in the same column correspond to the same fitting process with 3000 iterations. Upper limit for *trexp* parameter for each fitting process is indicated at the top, the obtained optimal value of this parameter was near the upper limit. Other parameters limits were identical (s. Table 3.3).
Panel A shows V_{max} for PFK and FBP, calculated from gene expression data with fold change amplification parametrised by the exponent *trexp*. *Panel B* shows the metabolic term g for PFK and FBP. *Panel C* shows experimental and fitted net flux. Calculated FBP flux is indicated additionally with negative sign. *Panel D* shows flux the difference analysis for the corresponding fit. Shown are the normalised contributions to flux difference from the 0% state from hierarchic (R_H), metabolic (R_M), and unknown (R_{unk}) source adding contributions from both PFK and FBP .
All calculations were performed sample wise, mean and standard deviation of the resulting distributions are shown, except for the measured flux for which both values per condition are shown. As transcription amplification, and hence hierarchic regulation is reduced, metabolic regulation is assigned a more pronounced role when fitting to measured net flux.

Flux dissection analysis, shown in panels D, Fig. 3.8, reveals that R_M values increase with lower $trexp_{max}$ values for all aerobic conditions. This indicates that the fit to data is achieved by using parameters values resulting in an increase in metabolic regulation which is indeed able to counterbalance the reduction of differences in V_{max} between conditions (implied by lower values of the parameter *trexp*). Higher single-sample variances may be interpreted as a direct effect of higher metabolic regulation, since this also amplifies the generally high variances of metabolite level sample values.

An interesting point is the sharp upregulation of the transcription *FBP1* from 2.8% to 21% condition connected to the question of a potentially active futile cycle between PFB and FBP. Under our assumption of 'concerted regulation' (s. Eq. 3.16), this suggests a non significant amount of active FBP enzymes. At the same time, the FBP inhibitor metabolite F26bP is predicted to exhibit very low levels. From a purely reaction-kinetics point of view, the assumption of significant flux through FBP in the 21% condition does offer a convenient explanation for the controversy presented by the sharp rise of the PFK substrate F6P from 2.8% to 21% condition (with roughly constant level of the its product F16bP), compared to the flux decrease between these conditions.

This, however, introduces an actual futile cycle in glycolysis. Since the flux through PFK equals the ATP consumption rate, the specific ATP-cost per converted G6P molecule is determined by the rates through PFK and FBP (we neglect the flux towards F26bP for this calculation) and can be written as

$$\frac{\frac{d}{dt}ATP_{\mathrm{G6P}\rightarrow\mathrm{F16bP}}}{V_{\mathrm{net}}} = \frac{V^{\mathrm{PFK}}}{V^{\mathrm{PFK}} - V^{\mathrm{FBP}}} \tag{3.36}$$

For negligible FBP rates, as generally assumed to occur in glucose consuming metabolic states, the ATP-cost per F6P molecule converted by PFK is practically one. Assuming a glucose consuming metabolic cell state with non-neglectable FBP-rate exists, this would increase the ATP cost.

During anaerobic metabolism of glucose, the activity of FBP must be neglectable, otherwise the low gain of only two ATP per glucose molecule would be drastically reduced by the futile cycle. However during aerobiosis ATP gain per glucose may reach well above 20 molecules per glucose (depending on the relative ratio of aerobic metabolism) some loss of ATP through simultaneous FBP activity might be an minor loss if it comes with advantages.

Such an advantage may potentially be the increased flexibility with regard to quick regulation of the net reaction rate, hence the flux ratio between glycolysis and branches from G6P via change in the FBP rate.

Sudden oxidative stress in certain phenotypes is an example for a scenario in which the ability of quick flux redirection towards the pentose-phosphate-pathway, and resulting increased NADPH production, may be of advantage [66]. On the other hand, flux redirection towards the main glycolytic pathway may be of advantage, if the sudden stop of oxygen provision, resulting in a sharp drop of ATP gain per consumed glucose. In some cases, the ability to switch to gluconeogenetic state on a faster timescale than *de-novo* FBP production can be started, may be an advantage. These scenarios are potentially connected to aerobiosis.

However, both PFK and FBP are allosterically affected by F26bP, AMP, and citrate (the two latter are not included in the model kinetics). Hence, the

above described hypothetical advantage requires the assumption that allosteric regulation of PFK alone is limited either in reaction time, or in extent. In this case regulation via an additional enzyme may offer advantage.

In agreement with the above considerations, expression of *FBP1* was practically zero in the anaerobic case and was only of significance in the two conditions with highest oxygen provision. Under the assumption of concerted hierarchic regulation i.e. that enzyme activities change (approximately) in accordance with transcription activity, the emerging picture is that a futile cycle between PFK and FBP is indeed active in one or more aerobic conditions, and consumes some of the generated ATP. However, even in the 21% oxygen condition when expression of PFK was lowest and that of FBP highest, signals for PFK were still four times higher.

3.4 Discussion

3.4.1 General considerations

In this chapter, we used data from different levels of physiology to study regulation of metabolism on the level of single enzymatic reactions. The data are based on glucose limited continuous yeast cultures with constant growth rate $D = 0.1/h$. The free experimental parameter was oxygen supply: cultures with 0% (anaerobic), 0.5%, 1%, 2.8% and 21% (fully aerobic) of oxygen volume part in the inlet gas were analysed, [97] [47] [69]. Particulars of the experiments and the datasets were discussed in Chapter 2.

In particular we asked the following questions:

- (i) To which extent can we link the three datasets by a kinetic equation model (and certain specified assumptions), in a way consistent with biological knowledge, at the level of a single enzymatic reaction.

  In our approach, this was judged by the extent to which measured flux changes could be reproduced by the reaction rate calculated from the kinetic equation model.

- (ii) Can we assign the sources of flux difference between conditions to the two basic levels of regulation termed 'hierarchic' and 'metabolic' after Westerhoff et al. [53].

  Since this analysis necessarily involved calculations based on enzyme kinetic models, it could only be tackled to the extent, to which these models were seen as satisfactory, in terms of (i).

The 'assumptions' in (i) included, as usual, the assumption that most quantities with no available information are constant or do not matter. Other assumptions, such as those connected to transcription, are described below.

The question posed in (i) generally corresponds to a certain stage in data-based modelling, and should be answered as 'satisfactory' before further conclusions are drawn from the modelling. In our case, goodness of reproduction of the flux change, hence the applicability of the model often varied between experimental conditions, when analysing one reaction. In cases when the answer

to (i) was a clear 'bad', we based our argumentation on data, rather than the model.

We note that a negative answer, i.e. the fact that flux change *could not* be reproduced in a certain case, may well be seen as information of interest, especially when the search range of the model was well defined, and it is analysed, the falseness of which model assumptions may be responsible for the failure, which was done with our best knowledge.

Answering (i) was attempted by using a fitting procedure. The models with best-fit parameter sets were studied further. Importantly, the fact that some parameters were not restricted further by the fitting process (than they were a priori by the search limits), did not alter the answer to (i), since, in the given context, this translated to whether or not it *is possible* to reproduce the flux data using a certain model which is *consistent* with biological knowledge.

In those cases when the answer is 'good', i.e. data reproduction is judged to be satisfactory, definiteness of parameters does not play a role for (i), as long as all parameters stay in a range consistent with biological knowledge.

In case when the answer is 'poor', i.e. data reproduction is judged to be poor, definiteness of parameters plays an even less important role for (i). In this case, parameter search ranges arguably become more important, since the negative answer is never absolute, and the information content is higher if larger variety of parameter values (or models) were tested (and documented).

For the question posed in (ii), we found that the quantities derived from the best-fit models were rather robust with regard to the fitting procedure. This implied that even if some of the parameters of the underlying models were not identified by the fitting process, all best-fit sets of parameters resulted in very similar answer. In our study, this included the transcription amplification parameter $trexp$, the calculated Vmax values, and reaction rates, and the flux dissection coefficients mentioned below.

Hence we concluded that parameter identifiability or even 'overfitting' did not impair our conclusions as long as all best-fit results lead to the same answer.

3.4.2 Methods

In line with the general purpose of this work, we attempted to elucidate the regulatory background underlying flux redistribution between experimental conditions. For a single enzymatic reaction this was translated to the question posed under (ii) above: can flux difference between conditions be dissected into parts assigned to 'hierarchic' and 'metabolic' regulation, respectively?

Earlier work of Westerhoff *et al.* [53] introduced the idea of regulation analysis ie. dissecting flux changes through enzymes into 'hierarchically' and 'metabolically' regulated parts (corresponding to the regulation coefficients ρ_h and ρ_m respectively, reproduced in Eq. 3.3). This was based on enzyme activity and flux data. They did not use metabolic quantification data, hence they could not use explicit reaction kinetic expressions to estimate reaction rates.

On the other hand, they utilised a rather general property of rate equations of catalysed reactions of the form of Eq. 3.1: the linearity of the relationship be-

tween the enzyme activity Vmax and the reaction rate. Combining this with the reasonable assumption that fold changes in measured enzyme capacity equal fold changes of *in vivo* Vmax values, they derived the statement that the deviation of measured fold change in flux through an enzyme from corresponding enzyme activity fold change is due to metabolic regulation, i.e. associated with the term g in Eq. 3.1. (After normalisation, this notion is expressed by the relationship $\rho_h + \rho_m = 1$ which was termed[6] 'summation theorem').

In a later work, [73, Rossel *et al.*], they classified reactions according to the relationship of the regulatory components between two conditions as 'cooperative' (metabolic and hierarchic regulation support each other), 'conservative' (metabolic regulation prevails over hierarchic regulation, the two acting antagonistically), or 'antagonistic, directed by Vmax' (hierarchic regulation prevails over metabolic regulation).

Since we used a corresponding metabolic quantification dataset, we could, in addition, utilise existing biochemical knowledge in form of reaction kinetic models to estimate reaction rates and compare them to measured flux values. This allowed us to dissect the measured flux change into a reproduced portion and that of unknown source (denoted by R_{unk}, quantified in Eq. 3.11).

However, since we use transcription data rather than enzyme activity data, in contrast to Westerhoff *et al.* this unpredicted portion of the flux change has a different interpretation, since it can not be simply assigned to metabolic or hierarchic regulation: it may either stem from post-transcriptional regulation, or from metabolic interactions not represented in the dataset or the model. A further complication was introduced in the cases when a higher number of genes are associated with the same reaction.

The reproduced portion of the flux was further dissected into metabolic and hierarchic parts, with the corresponding coefficients, R_M, and R_H respectively, defined in Eq. 3.8.

While the biological notion for this latter dissection was closely related to that of Westerhoff *et al.*, we found that their mathematical formulation did not entirely suit our needs, since it suffers from a number of limitations. Notably, we could not apply the formula in the case when two enzymes are associated with the same reaction, as this is the case with the PFK-FBP enzyme pair. Our mathematical formulation for the flux dissection, is based on flux difference, rather than flux fold change, resulting in the above coefficients having a linear relationship with flux change, a property which eases their interpretation in case more enzymes are involved.

Since transcription data change were taken as a proxy to enzyme activity, a concept of linking transcription with enzyme activity was needed, if only to consistently formulate and biologically interpret the assumptions employed.

For this purpose, a simple model of hierarchic regulation was developed within which we could formally formulate the assumption of 'concerted hierarchic regulation' in the form of Eq. 3.16. The resulting one-parameter formula, Eq. 3.18, links fold change between steady state levels of mRNA and that of protein activity. The single parameter *trexp* was associated with 'transcription

[6] Somewhat contrary to mathematical tradition.

amplification' and, in case the assumption of concerted regulation holds, was shown to correspond to the number of hierarchic regulatory levels. Hence, Eq. 3.18 was introduced into the enzyme kinetics models with *trexp* as a further parameter using search range limits [0-5], the upper limit being the approximate number of known regulatory levels.

While mRNA and protein level correlation has been reported moderate (e.g. [45], [5], [27]), a newer study has reported high correlation values after up to four hours after cells were subjected to osmotic shock by NaCl [54, Lee *et al.*]. Interestingly, strong average correlation was only reported for upregulated genes.

In line with other studies, Lee *et al.* also reported that, on average, protein level fold changes were slightly lower than mRNA level fold change. The average 1.4 fold mRNA upregulation corresponded to an average protein level change of 1.3 fold; this corresponds to *trexp* value of around 0.8 (given values are approximate).

However, in our case, best-fit results for the transcription amplification parameter *trexp* were between 3-5, significantly higher than the current consensus. This deviation may stem from a number of facts. As discussed in Section 3.2.2, the transcription dataset used in this work was generated using Affymetrix genechips for which linearity of the relationship between signal and mRNA levels is not guaranteed.

The above high-throughput studies aimed at elucidating general tendencies rather than single-gene cases. In addition, great variance in the levels of correspondence between mRNA and protein fold change was reported between average values for functional groups of genes. Usually, such studies measure protein level, while we used the notion of enzyme activity, corresponding to levels of active protein, which, in some cases is different from total enzyme level.

There is also a significant difference in experimental conditions regarding time scale. Genome-scale comparison studies are typically based on cells harvested after a few hours after perturbation (e.g. exposing the culture to a certain stress condition). This is significantly different from the 40 hours (six generations) which the cells spent in a (nearly) identical environment before samples were harvested from our continuous chemostat cultures. Hence, different relationship between mRNA and protein fold changes under such conditions is, although not expected, not necessarily inconsistent with the above results.

Nonetheless, transcription data are not in all cases well suited to answer questions on enzyme activity. On the other hand, in cases, when the enzyme of interest is notoriously difficult to characterise experimentally, transcription data may offer more reliable information. Difficulty of measuring PFK activity *in vitro* might have been a potential reason for the intriguing failure of reproducing PFK flux change in [9, van den Brink]

3.4.3 Results

Overall, with regard to the question posed under (i), we found that flux change through the studied enzymes could be reproduced to largely varying extent. Accordingly, in some cases, flux dissection into hierarchic and metabolic parts could be performed only in part.

Surprisingly, flux changes through the enzyme phosphogluco-isomerase (PGI) could be reproduced to the lowest degree, although this enzyme is not renown for the complexity of its regulation. Hence, a potential reason of this may be the difficulty of measuring reactant levels as these are especially sensible to quenching delay variances due to their closeness to the extracellular environment in the reaction chain.

While failure in the 1% oxygen condition may potentially be accounted to false metabolite levels, flux decrease towards fully aerobic (21%) condition is incompatible with both metabolite and transcription data. Assumption of systematic error in the metabolic dataset seems the most probable cause at this point. This would imply antagonistic regulation from the hierarchic and the metabolic levels between the 2.8% and the 21% conditions, with metabolic regulation getting the overhand, placing PGI in the 'conservative' category between these conditions, according to [73, Rossel *et al.*].

Flux changes through the pyruvate dehydrogenase complex, could be reproduced slightly better: even though absolute flux values were reproduced with the exception of the 2.8% condition, the basic increasing tendency of the measured flux towards higher oxygen levels was only roughly reproduced (mainly due to a transcription fold change between 0% and 0.5% conditions), but not in finer details. In contrast to the case discussed above, this was not entirely unexpected, given the elaborate phosphorylation-based regulation mechanism of this complex which could not be modelled in a single reaction model, due to lack of data, especially on Acety-CoA, the main metabolic effector of the regulatory kinase/phosphatase pair of the complex. Using citrate as a proxy for its concentrations resulted in only minor improvements.

With respect to regulation, it could be seen from the pyruvate level pattern that the reaction was not regulated by substrate concentration change. Flux change from 0% to 0.5% level condition could be reproduced. For this change, cooperative regulation with both metabolic (due to citrate level change), and hierarchic components could be determined, However, lack of flux change reproduction of the model did not allow to make such conclusions about the other conditions.

It remains an open question, to which extent substrate level changes acting via the kinase/phosphatase pair play a regulatory role. This is would be intuitive insofar that a metabolic regulation-based mechanism would allow the cell to switch between respiratory and fermentative metabolism quickly.

Notably, the phosphorylation-based regulatory mechanism of this reaction seem not to fit well conceptually into the categories 'metabolic' and 'hierarchic' created by Westerhoff *et al.*, since the enzyme activity (Vmax) of the PDH-complex itself is regulated by a process (phosphorylation) which is significantly effected by both hierarchic regulation of the kinase-phosphatase-pair, as well as levels of the reaction product.

For the TCA-cycle reaction catalysed by fumarase, flux changes were mostly well reproduced by the calculated reaction rates. Notably, metabolic regulation acted mostly (slightly) antagonistic to flux change, the latter being 'forced' by hierarchic regulation. Hence, with respect to the conditions under study, this enzyme falls into the category 'conservative' according to [73, Rossel *et al.*].

Interestingly, the flux-peak at 2.8% condition with flux decrease at 21% could not be reproduced from the data, neither for the pyruvate dehydrogenase complex, nor for the much simpler case of fumarase. While in the first case there may be potential explanations, the latter presents a more intriguing surprise, given that the otherwise intuitive data seem to clearly imply a flux change - but in the wrong direction. Setting aside the possibility of a systematic measurement error, this seems to be the clearest case when failure of the assumption of 'concerted hierarchic regulation', relating transcription data to enzyme activity, is indicated. However, this effect could also be a manifestation of the non-triviality of fumarase kinetics, even though current knowledge places our experiments into the Michaelis-Menten regime of this enzyme [57].

In any case, this indicates a still substantial difference in metabolic states between these two anaerobic conditions, even though reported flux changes are relatively small.

Somewhat surprisingly, data on the complex regulatory network of the reaction associated with phosphofructokinase (PFK) and fructose-bisphosphatase (FBP) was rather successful. This is strongly connected to the fact that most transcription and metabolic data were in accordance with the reported flux change. However, the slightly higher number of parameters, most notably allowing to adjust for the absolute activity of the gluconeogenetic enzyme FBP, was another factor. Nevertheless, without consistency of data and model, such high extent of flux reproduction would not be possible.

An interesting feature of the described model of PFK-regulatory network was discovered: reducing transcription amplification resulted in better reproduction of the decreasing tendency of net flux towards higher oxygen levels, while the modification lead an increased role of metabolic regulation in the reproduced rate change, and higher FBP activity in the high-oxygen provision conditions.

This might be interpreted as an indication of the ability of this regulatory network - and the measured metabolic levels - to 'replace' hierarchic regulation by metabolic regulation. This may be an advantage when reaction time is too short for hierarchic regulation to react. However, this property of the PFK-FBP model is more fittingly studied in the context of a larger dynamic model, allowing time-dependent simulation of associated scenarios.

High FBP expression levels and improved fit to flux data may be interpreted to suggest significant presence of FBP activity, even though this would lead to futile cycling of F6P and F16bP, and hence, to ATP loss.

While theoretical investigations on this cycle go back to [79, Sel'kov], current consensus leans towards the opinion that, even though this futile cycle may be used in some organisms for heat generation ([84], [44]), it is generally avoided by allosteric regulation. It was demonstrated that FBP, if artificially expresssed in yeast, can be depressed via allosteric control, such that futile cycling is below 2% [11]. Still, to our knowledge no final consensus is available on the question to which extent the catalytic activity of translated FBP is downregulated in all glucose-consumption growth conditions in yeast [75, U. Sauer, pers. comm.]. This is related to the high effort needed for direct experimental evidence in this question, which can not be provided by the presented flux measurements, since only the net flux $V_{net} = V_{\mathrm{PFK}} - V_{\mathrm{FBP}}$ is measured.

However, our interpretation of the presented flux, metabolic, and transcription

datasets provides some indications of a scenario that these antagonistic enzymes worked at the same time under aerobic conditions. We hypothesised that this might indeed the case if the relatively low ATP loss (under aerobic conditions) improves flexibility to react to external perturbations quicker or to a higher extend. We note that, for budding yeast, chemostat growth conditions are not regarded as resembling that in nature.

A potential explanation may be that, in spite of mixing, single cells do experience some periodic glucose level changes in a chemostat, which, due to the low residual glucose concentrations inherent to the system, are sufficient induce periodic allosteric upregulation of some gluconeogenetic enzymes.

The above considerations are intimately linked to the fact that the data were generated from steady state samples. A picture according to which perturbations are first dealt with by metabolic regulation and - if persistent - hierarchic regulation takes over, is consistent with these considerations. Especially the results for the PFK-FBP system may be interpreted as to suggest that, in continuous cultures, the basic metabolic cell state is already set at transcription level while further details are regulated at metabolic (and possibly) levels.

The above considerations may be complemented by the double perturbation experiment by [9, van den Brink *et al.*]. Starting from a continuous culture very similar to our glucose limited 21% oxygen condition, they introduced a persistent perturbation by turning off oxygen and significantly increasing glucose levels. Interestingly, using the same PFK kinetics we used in Eq. 3.25 (without FBP), and measurements of F6P, F16bP and F26bP levels, as well as of enzyme activities (which was, however, reported to stay constant), the sudden PFK flux increase could not be reproduced at all in [9]. It remains to be seen, if inclusion of FBP into the model offers a different answer.

On the whole, flux changes which could be reproduced by the modell, tended to be sourced back to cooperative regulation within the models, with the notable exception of fumarase when the two regulatory levels acted antagonistically. This is in line with van den Brink *et al.* (s. Figure 6 in [9]) who reported approximately 20%-50% of the flux change to come from enzyme activity change for the enzymes phosphoglucoisomerase, pyruvate kinase, and pyruvate decarboxylase, 120 minutes after the perturbations (even though this measure is not directly comparable to the R_H values used above). In line with the considerations above, increasing replacement of metabolic regulation by enzyme activity increase was reported for these three enzymes.

In the next chapter, we will attempt to make use of the above results and put the above notions to test in a dynamic model of central carbon metabolism.

Chapter 4

Using the building blocks: modelling central carbon metabolism

4.1 Introduction

4.1.1 A system wide approach to central carbon metabolism

In this chapter, we change the focus of our investigation to a larger scale: we will consider pathways making up significant part of central carbon metabolism. We will aim to investigate, to what extent a kinetic model - containing a considerable amount of current biochemical knowledge - is consistent with the datasets presented in Chapter 2.

The model which has been constructed for this investigation is a modification and extension of the kinetic glycolysis model by Teusink *et al.* [87]. The model was modified such that, in addition to anaerobic, also respiratory energy metabolism is represented, albeit in a highly simplified way. This makes it possible to represent metabolism in anaerobic as well as aerobic yeast cultures.

In order to provide flux distributions which are consistent with the model's stoichiometry, the constrained flux balance analysis procedure used in Jouhten *et al.* [47] was partly repeated, using a modified stoichiometric network.

The result may be characterised as a kinetic model which is stoichiometrically consistent with a larger stoichiometric network. Beyond the immediate aim of this study, it is presented in the hope to be useful for further studies as well as a basis for further extensions.

Modelling central carbon metabolism in detail is somewhat of an ambitious attempt. The pathway is naturally subject to constraints stemming from its stoichiometry - involving several cofactors - and the laws of chemical kinetics. Additionally, the glycolytic pathway is known to include a number of allosteric regulatory links, which, together with other types of regulation, influence its behaviour. Fig. 4.1 shows a selection of the regulatory links known to mediate information.

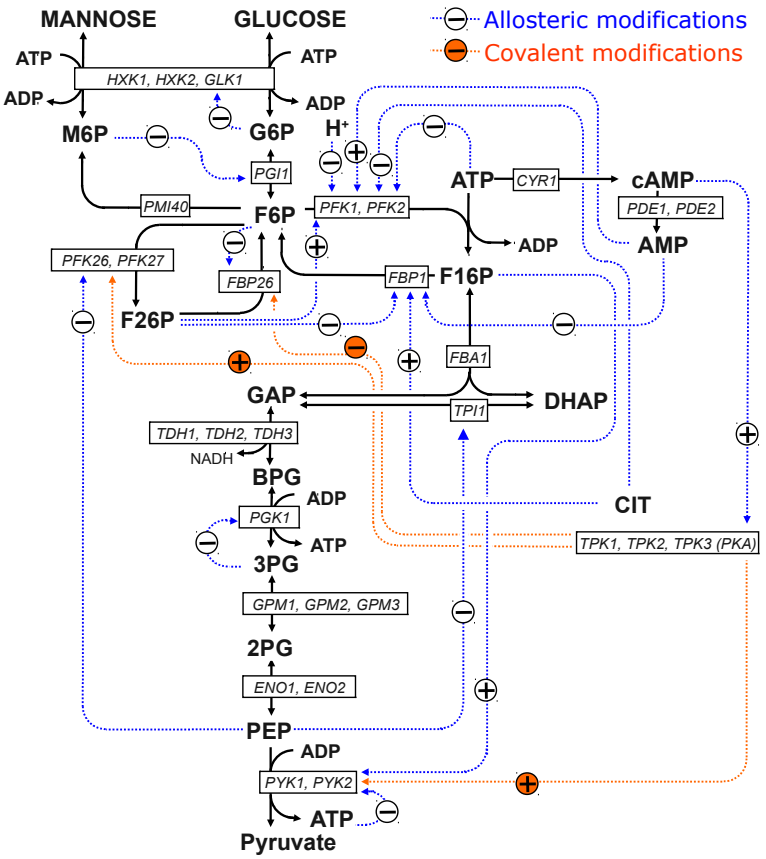

Figure 4.1: Scheme of the glycolysis with a number of known regulatory links. Based on a figure in [64].

Before we start our investigation, we give a brief review of the special form of ordinary differential equations, used for modelling biochemical reaction networks.

4.1.2 The formalism: describing chemical reaction networks

Spatially homogeneous ('well stirred') time evolution of the concentration of n_S chemical species due to n_V chemical reactions is often described by a system of ordinary differential equations of the following form

$$\frac{d}{dt}S_s = \sum_{v=1}^{n_V} N_{sv} V_v(S,P) \quad s = 1...n_S \tag{4.1}$$

where the variable vector[1] $S = \{S_s, s = 1...n_S\}$ denotes the list of concentrations, and the kinetic rate expressions $V_v(S,P)$, $(n = 1...n_V)$ are associated

[1] We denote indexed quantities with upper case letters and furnish them with lower case indices. If not stated otherwise, the name of such quantity without an index will denote the whole array, e.g. $V_v(S_s, P_p)$ may be shortened to $V(S,P)$.

with reaction rates. The parameter vector $P = P_p$, $p = 1...n_P$, contains quantities regarded to be constant for the time scale of interest. This often includes 'external' species, assumed to be virtually constant due to their supply from a large reservoir.

The stoichiometric matrix N_{sv} encodes the information which reaction transforms which species into each other. N is only determined up to a constant which is traditionally chosen such that the entries of the matrix - the stoichiometric coefficients - attain the smallest possible integer value. Such a choice exists for networks representing typical chemical reactions between molecules of a well defined numbers of constituents, however, networks including other types of processes - such as the electron transport chain appearing in this chapter - may be described by a stoichiometric matrix containing non-integer (or even irrational) coefficients. Stoichiometric coefficients are assumed to be constant in this work.

The stoichiometry of the system Eq. 4.1 is defined by the stoichiometric matrix N together with the list of reactions considered irreversible, i.e. subject to a constraint on their sign. If only this information is given without Eq. 4.1, we will speak of a stoichiometric network or stoichiometric model of the represented reaction system.

For chemical reaction networks, the kinetic rate expressions are of equal importance as the concentrations S, the actual variables of the system. For biochemical systems, they are often of the form

$$V_v(S, P) = V_{\max}\, g(S, P) \tag{4.2}$$

This is the form of kinetics we have discussed in section 3.2.1. We will assume that each rate expression contains at least one parameter. The above also includes constant rate expressions of the form $V(S, P) \equiv V_{\max}$ (i.e. $g(S, P) \equiv 1$) often used to represent external reactions.

The differential equation Eq. 4.1 implies a time evolution of the concentrations and the rates which therefore may be written as $S(t)$ and $V(S(t), P)$, respectively. We will assume that, for any admissible values for the vectors (S, P), a unique solution of the system Eq. 4.1 exists. This is guaranteed if the right hand side of Eq. 4.1 satisfies a condition known as Lipschitz continuity which is normally the case when describing biochemical reaction systems.

Eq. 4.1 possesses a steady state, if there exists a concentration vector S^* for which the vector $V(S^*, P)$ lies in the kernel of the matrix N, i.e. it fulfils the steady state condition

$$0 = \sum_v N_{sv} V_v(S^*, P) \tag{4.3}$$

The existence of steady states is not necessarily the case. When using kinetic expressions of the form Eq. 4.2, parameter sets often exist which can be interpreted as describing a 'blocked' reaction in a pathway resulting in the *ad infinitum* accumulation of one or more species.

Some systems exhibit oscillatory behaviour for certain sets of parameters and initial conditions. While this is relevant to the modelling of glycolysis, we will not study such behaviour in this work.

Steady state reaction rates in a chemical reaction network are called *fluxes* and denoted by J

$$J(S^*, P) := V(S^*, P). \tag{4.4}$$

It should be pointed out that $J(S^*, P)$, a flux distribution of the reaction system, is a different mathematical object from the rate $V(S, P)$: while a rate can be assigned to any vector of concentrations and parameters, i.e. V is a function on the space of all feasible (S, P), the flux is only defined for those combinations satisfying Eq. 4.3. In other words, the function $J(S^*, P)$ is only defined on the zero-level-set[2] of the function defined by the right side of Eq. 4.1.

The relationship between stoichiometry and kinetics is not trivial. It is relevant for the present chapter to realise, that the existence of a flux distribution - a vector within the kernel of N - does not imply the existence of a corresponding steady state of Eq. 4.1.

A NOTE OF CAUTION, REGARDING the term 'steady state', has to be issued to the reader (c.f. Chapter 2). Apart from the above use in dynamical systems theory, it is used in biology to describe cultures with no temporal change of the measured variables, such as continuous chemostat cultures (even though, such cultures are sometimes associated with 'exponential growth', since they imply a non-zero growth rate).

But metabolism of cells in a 'transitional' culture exhibiting temporal change - such as the growth rate in a typical batch culture - may still be best described using quasi-steady states of the metabolic reaction network. This results from the apparently rather clear timescale separation between dynamics purely on the metabolite level - known to attain a steady state within a 0.5 - 5 minutes even after a major perturbation, c.f. [91], [90], [93], [12] - and that on the enzyme level. For the latter, typical changes seem to be slow enough, [9], for the assumption to be made that enzyme level changes leave metabolite levels very close to their steady state values. In the conceptual framework centred around Eq. 4.1, the above assumption states that the enzyme level changes cause the system to go through a parametrised set of steady states.

Hence, the above assumption implies that measurements performed after the first few minutes after a perturbation do not to offer the possibility to observe the dynamics of the metabolic regulatory level itself as directly as in [90] , even if they are rightly called 'transitional' from the observational perspective.

Several texts across a variety of fields discuss systems of the form of Eq. 4.1 with special emphasis on describing biochemical or metabolic reaction networks, including [36, Heinrich and Schuster] [24, Érdi and Tóth], [26, Fell], [50, Klipp *et al.*] [85, Stephanopoulos *et al.*] [59, Nielsen *et al.*].

4.1.3 Consistency with an existing kinetic model

Teusink *et al.* published a detailed kinetic model of yeast glycolysis [87], based on measurements on non-growing cells in anaerobic batch cultures with glucose concentration starting from 100 mM. Enzyme activities and many enzyme kinetic parameters were measured *in vitro*, some parameters were taken from pre-existing

[2] Under certain requirements which are mostly ignored in introductory texts, the level set of Eq. 4.1 forms a smooth 'surface' or differentiable manifold. For example the condition that the right hand side is differentiable and its gradient non-zero on the whole set guaranties the level set to be a smooth hypersurface in the space (S, P)

literature. For the convenience of the reader, we collected known misprints in [87] in Appendix B.

For investigating the consistency of the model with the datasets presented in Chapter 2, we had to take into account the differences between experimental conditions (c.f. section 2.3). In contrast to Teusink's, our data were based on cultures in glucose limited conditions implying very low external glucose concentrations. Another notable difference was the non-zero specific growth rate of 10% per hour in biomass in the chemostat cultures.

The difference in the culture conditions resulted in different external environments for the cells in terms of external metabolite concentrations. Extracellular ethanol concentration in the chemostat cultures was 75 mM, somewhat higher than the 50 mM in Teusink *et al.*, however the glycerol concentration of 9 mM was more than an order of magnitude higher than the 0.15 mM in [87].

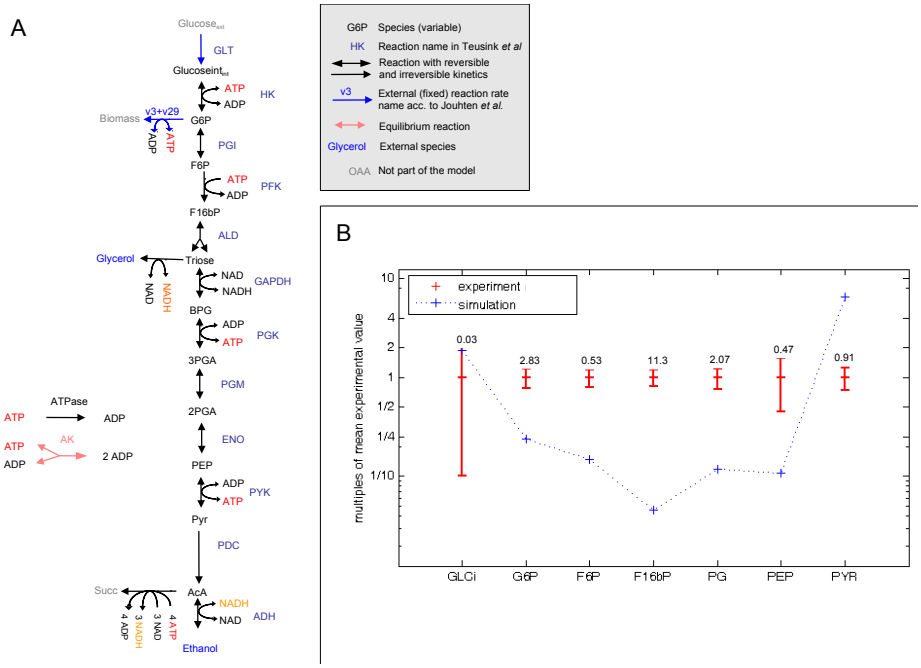

Figure 4.2: Comparison of chemostat data from [47], [97] with predictions of the Teusink model. *Panel A*: Scheme of the model. See text for differences from [87]. The reactions TPI and AK are assumed to be in equilibrium and not represented by kinetic expressions. The metabolites GAP and DHAP are represented by a single variable called triose. *Panel B*: Comparison of the simulated steady state concentrations with experimental values corresponding to anaerobic chemostat cultures. Each simulated value is plotted as multiple of the corresponding mean experimental value which is shown as unity. Errorbars show standard deviations of the experimental data. Mean experimental values are shown in mM above the error bars.

Consistency of the presented dataset with this kinetic model was investigated by asking, to what extent the model reproduces the intermediate metabolite concentrations with the external conditions set to match those in our anaerobic chemostat cultures. In order to be able to answer this question, a minimal set of modifications had to be undertaken at the model.

Since the glucose concentration in the chemostat culture was below detection limit, this quantity could not directly be used. Instead, the flux of the glucose transport reaction was set to a constant value of the measured glucose flux into the cells in the chemostat cultures, 68 mM/min. This value is somewhat lower than both the simulated (88 mM/min) or the experimental value (108 mM/min) in [87].

In addition to the above alterations in external conditions, the branches towards glycogen and trehalose in the Teusink model were merged into a single branch for which the corresponding flux was set to the sum of the fluxes $v3$ and $v29$ in our flux dataset (c.f. Fig. 2.2).

The scheme of this intermediate model and a comparison of the simulated steady state internal metabolite values with the experimental data is shown in Fig. 4.2. The simulated concentration values lie well outside the experimental error, typically differing more than 4 fold from the mean experimental values.

For the aerobic (20.9% oxygen) condition, using our dataset to define nominal external concentrations and fluxes in an analogous way, did not produce a steady state.

Failure of the model in reproducing our chemostat culture data is not surprising since the original model - including its parameter set - was constructed to describe data obtained from a different cell line subjected to external conditions different from ours. Most notably, our dataset is based on glucose limited chemostat cultures, not on batch cultures. As discussed in Chapter 2 in detail, while the cells are objected to an initially high, albeit sinking glucose concentration in the latter, cells in a glucose limited chemostat culture are expected to be optimised for achieving the pre-defined growth rate (dilution rate D) with the lowest external glucose concentration.

These differences correspond to different internal states of the cells, between batch and chemostat cultures. Regarding the glycolytic pathway represented in the model, we expect that at least enzyme activities - represented by maximum rate parameters ($V_{\max}$) in the kinetic model - differ between cells from these different cultivations.

Before we set out to explore the capabilities of the model to reproduce the data with modified maximum rates for its enzymes, it is necessary to modify it to include respiratory metabolism, and to consider the inclusion of our experimental data in more detail.

4.2 Model construction

4.2.1 Using flux data in the kinetic model

Using the flux dataset presented in Chapter 2 in the kinetic model to be presented proved a non-trivial task. If a kinetic model's steady state is to reproduce a given full set of fluxes - in our case one associated with measurement data - it is necessary that this flux distribution is admitted by the stoichiometry of the model as a balanced flux distribution, independently of the used set of reaction kinetics.

While this is a necessary condition, it does not imply that the kinetic equations are such that the required flux distribution can be produced with a (feasible) set of metabolite concentrations. This may be achieved by changing parameters of

the kinetic equations as long as stoichiometric feasibility of the objective flux is satisfied. Ignoring this requirement may lead to frustratingly fruitless efforts of parameter search to produce the desired flux distribution. This situation may arise in practice, since it is often not intuitively clear - by visual examination of the diagram representing the stoichiometry - whether a model admits a certain flux distribution. Especially reaction networks including metabolic cofactors (such as adenosine nucleotides) are hard to judge by visual inspection, since these co-metabolites cause the reaction network to be highly interconnected.

As described in Chapter 2, the flux dataset we intended to use resulted from a constrained flux balance analysis procedure which itself used a stoichiometric network (c.f. Fig. 2.2) as input. The aim of this procedure was to find a flux distribution which satisfies certain constrains - stemming from experimental data - within the space of balanced fluxes of the stoichiometric network encoding pre-existing biochemical knowledge.

The stoichiometric network used in [47] was not compatible with our intended kinetic model, since it did not include the nicotinamide cofactors. Indeed, this flux distribution does not belong to the space of balanced fluxes of our kinetic model, consequently, trying to achieve a flux distribution directly derived from the one in Fig. 2.2 failed.

To produce a flux distribution that is balanced for our model, but at the same time consistent with the measurement data underlying the flux data in [47], the constrained flux balance analysis procedure had to be repeated using a modified stoichiometric network instead of the one used in the publication, c.f. Fig. 2.2. This was constructed to be similar to the original one, but stoichiometrically consistent with our kinetic model, i.e. possessing the same set of balanced fluxes.

4.2.2 A stoichiometric network with metabolic cofactors and respiratory chain

For the above reasons, and in contrast to the original network in [47], the modified stoichiometric network included the co-metabolite species ATP/ADP and NAD/NADH at various reactions.

More marked modifications were required in order to describe the aerobic modes of metabolism, in order to allow a simplified description of both anaerobic and aerobic modes with a consistent stoichiometry. To this end, a simplified representation of the respiratory chain was included, coupling oxygen uptake rate with the turnover rate of the redox cofactor pair NADH/NAD.

The modified stoichiometric network is shown in Fig. 4.4; Table 4.1 shows the list of reactions and metabolites defining this network. In the calculations explained in the next section, we used the representation of the network by the stoichiometric matrix as shown in Appendix B, Table B.1. The considerations leading to this network included the following.

The TCA cycle has different stoichiometric modes for anaerobic and aerobic metabolism. In anaerobiosis, the full cycle is active and produces 3 NADH which are being reused in the respiratory chain as shown in Fig. 4.3. For each NADH, this chain of redox reactions catalysed by a number of multi-enzyme complexes pumps 6 protons[3] from the inner mitochondrial matrix into

[3] In most eukaryotes the full chain pumps 10 protons per NADH through the membrane,

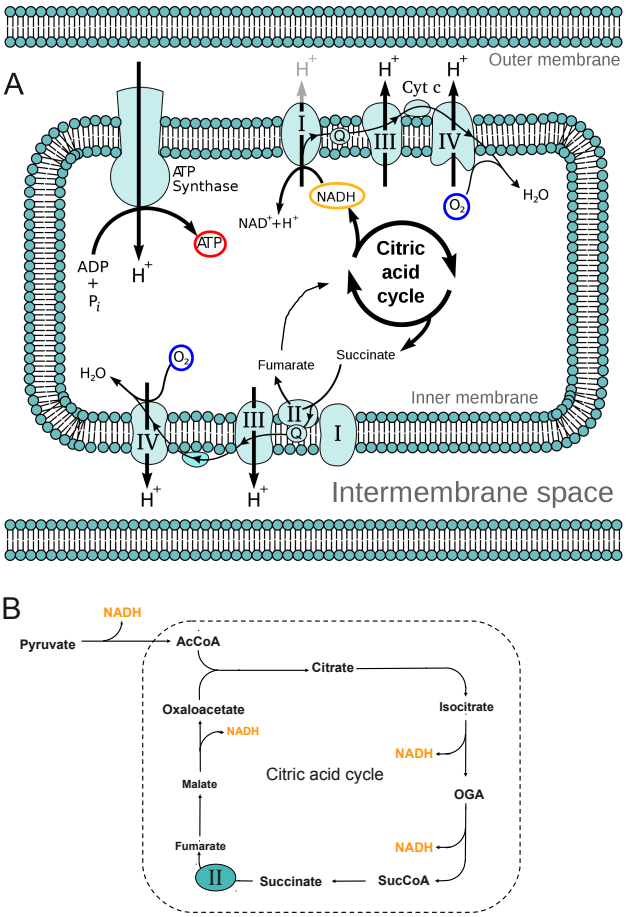

Figure 4.3: Schemes of the citric acid cycle and the respiratory scheme highlighting the stoichiometric coupling of oxygen, NADH, and ATP in a qualitative manner. Colour code for these species corresponds to that in Fig. 4.4.
Panel A: Respiratory chain. Indicated in the upper part of the inner membrane are the full electron transport chain starting from NADH oxidation, involving complexes I, II, and IV, as well as ATP synthase; these processes are represented as Reaction *v53* in the stoichiometric network and as Reaction *vRESP* in the kinetic model. Indicated in the upper part of the inner membrane is the electron transport chain, starting from succinate oxidised by complex II, which is incorporated in reaction *v13* in the stoichiometric network and *vRESP2* in the kinetic model.
In *S. cerevisiae*, NADH dehydrogenase assumes the role of complex I, however this enzyme does not pump protons into the intermembrane space (greyed out arrow) which was taken account in the stoichiometry of the models. Modified from a figure by Tim Vickers in Wikimedia Commons.
Panel B: Citric acid cycle. Complex II/succinate dehydrogenase is part of both the cycle and the electron transport chain. OGA: 2-oxoglutaric acid. Modified from a figure in [64].

however, *S. cerevisiae* possesses a simpler enzyme, NADH-dehydrogenase in place of complex I which does not possess proton pumping capability. Hence, for species with complex I, the stoichiometric constant for ATP in Equations 4.5 and 4.8 should be multiplied by the factor $\frac{10}{6}$.

the intermembrane space, using the free energy of the electron transfer onto oxygen - the terminal electron acceptor - one atom of which is reduced in complex IV [39],[95]. The membrane potential thus built up between the two sides of the inner mitochondrial membrane is used by ATP synthase - the amazing 'proton-mill', capable of regenerating an ADP molecule to ATP, for every $3\frac{1}{3}$ protons returning to the inner mitochondrial matrix through it.

Focusing on the overall stoichiometry of only a few species of interest, and rounding $3\frac{1}{3}$ to 3, we can write[3]

$$NADH + \frac{1}{2}O_2 + 2\,ADP \rightarrow NAD + 2\,ATP \tag{4.5}$$

Furthermore, electrons already enter the electron transport chain during the TCA-cycle. Complex II (or succinate dehydrogenase), catalysing the redox reaction

$$\text{succinate} + Q \rightarrow \text{fumarate} + QH_2 \tag{4.6}$$

links the stoichiometry of TCA-cycle flux and the oxygen uptake rate, since the two electrons transferred from succinate to ubiquinone (Q) enter the electron transport chain, to be finally transferred to oxygen. Again, focusing only on the species of interest, the analogous balance equation for the overall reaction is

$$\text{succinate} + \frac{1}{2}O_2 + 2\,ADP \rightarrow \text{fumarate} + 2\,ATP \tag{4.7}$$

However, this net ATP production assumes that each proton pumped into the intermembrane space will be used to regenerate ADP via ATP synthase. In reality, due to proton leakage, the total stoichiometric constant for ATP is lower, and its value may vary considerably between conditions, species, or tissues of multicellular organisms.

Hence, the above equations may be written in the following form[3]

$$NADH + \frac{1}{2}O_2 + R_{P/O}\,ADP \rightarrow NAD + R_{P/O}\,ATP \tag{4.8}$$

$$\text{succinate} + \frac{1}{2}O_2 + R_{P/O}\,ADP \rightarrow \text{fumarate} + R_{P/O}\,ATP \tag{4.9}$$

where $R_{P/O}$ denotes the so-called P/O-ratio, a phenomenological parameter often determined experimentally, linking the stoichiometry of ADP regeneration and oxygen consumption. For the cell cultures under study, this value was determined to be close to unity for all oxygenation conditions, around one third of the theoretical maximum value. Hence, in the construction in the stoichiometric network, we assumed:

$$R_{P/O} = 1 \tag{4.10}$$

The process described by Eq. 4.8 is represented as Reaction *v53* in the stoichiometric network. Reaction *v13* represents the process described by Eq. 4.9 lumped together with the last part of the TCA cycle after complex II, hence the production of 2 NADH of the latter appears in the balance of Reaction *v13*. While these reactions are reported to be inactive in anaerobic cultures in [47], the part of TCA cycle before complex II exhibits a small flux even in these cultures, presumably in order to meet demand for certain biomass precursors.

It was assumed that reactions in the stoichiometric network do not change the sum of ADP and ATP, i.e. consumption or production of ATP implies respective production and consumption of the same amount of ADP. This made the representation of one of them superfluous, hence only ATP appears explicitly in our network. Analogously, the redox-cofactor pair NAD/NADH is only represented via NADH.

The glyoxylate shunt is not included in the original network (Fig. 2.2) nor in our modified network, since this pathway was not active in the experiments [47].

v1	Glc + ATP	→	G6P + ADP
v2 (∗)	G6P	→	F6P
v3	G6P	→	R5P + CO_2
v4 (∗)	F6P + ATP	→	2 Trio
v5	2 R5P	→	S7P + Trio
v6	R5P + E4P	→	F6P + Trio
v7	Trio + S7P	→	F6P + E4P
v8 (∗)	Trio	→	Pep + ATP + NADH
v9 (∗)	Pep	→	ATP + Pyr
v10 (∗)	Pyr	→	$AcCoA_{mit}$ + NADH + CO_2
v11 (∗)	$AcCoA_{mit}$ + Oaa_{mit}	→	citrate
v12 (∗)	citrate	→	NADH + CO_2 + Oga
v13 (∗)	$\frac{1}{2}O_2$ + Oga	→	Oaa_{mit} + CO_2 + 2 NADH + $R_{p/o}$ ATP
v14	Oaa_{mit}	→	Pyr + CO_2
v15	ATP + Oaa_{cyt}	→	Pep + CO_2
v16 (∗)	YR + ATP + CO_2	→	Oaa_{cyt}
v17	acetate + ATP	→	$AcCoA_{cyt}$
v18	AcA	→	acetate + NADH
v19 (∗)	AcA + NADH	→	Ethanol
v20 (∗)	Trio + NADH	→	glycerol
v21	Oaa_{cyt}	→	Oaa_{mit}
v22	Oaa_{mit}	→	Oaa_{cyt}
v24 (∗)	Pyr	→	AcA + CO_2
v29 (∗)	G6P +ATP	→	Ø
v30	R5P	→	Ø
v31	E4P	→	Ø
v32 (∗)	Trio	→	Ø
v33 (∗)	Pep	→	Ø
v34	Oaa_{cyt}	→	Ø
v35	$AcCoA_{cyt}$	→	Ø
v38 (∗)	Oga	→	Ø
v40 (∗)	F6P	→	Ø
v51	Ø	→	O_2
v52	CO_2	→	Ø
v53 (∗)	NADH + $\frac{1}{2}O_2$	→	$R_{p/o}$ ATP
v54 (∗)	ATP	→	Ø

c1 (∗)	G6P
c2 (∗)	F6P
c3 (∗)	Trio (GAP+DHAP)
c4 (∗)	PEP
c5 (∗)	Pyr
c6 (∗)	AcA
c7 (∗)	OGA
c8	S7P
c9	E4P
c10	acetate
c11	Acetyl-Coa_{mit}
c12	Acetyl-Coa_{cyt}
c13	Oaa_{mit}
c14	Oaa_{cyt}
c15	cytrate
c16	CO_2
c17 (∗)	ATP
c18 (∗)	NADH
c19 (∗)	O_2
c20 (∗)	R5P

Table 4.1: Lists of formal reactions and species of the stoichiometric network. For easier comparison with the network presented in [47, Jouhten *et al.*], reaction name IDs present in both networks were kept identical; reactions not present in [47] were given ID numbers above 50. Subscripts *cyt* and *mit* denote cytosolic and mitochondrial species, respectively. Reactions and species incorporated in the our kinetic model are denoted by (∗). For the usage for species possessing mitochondrial complex I, see Footnote 3.

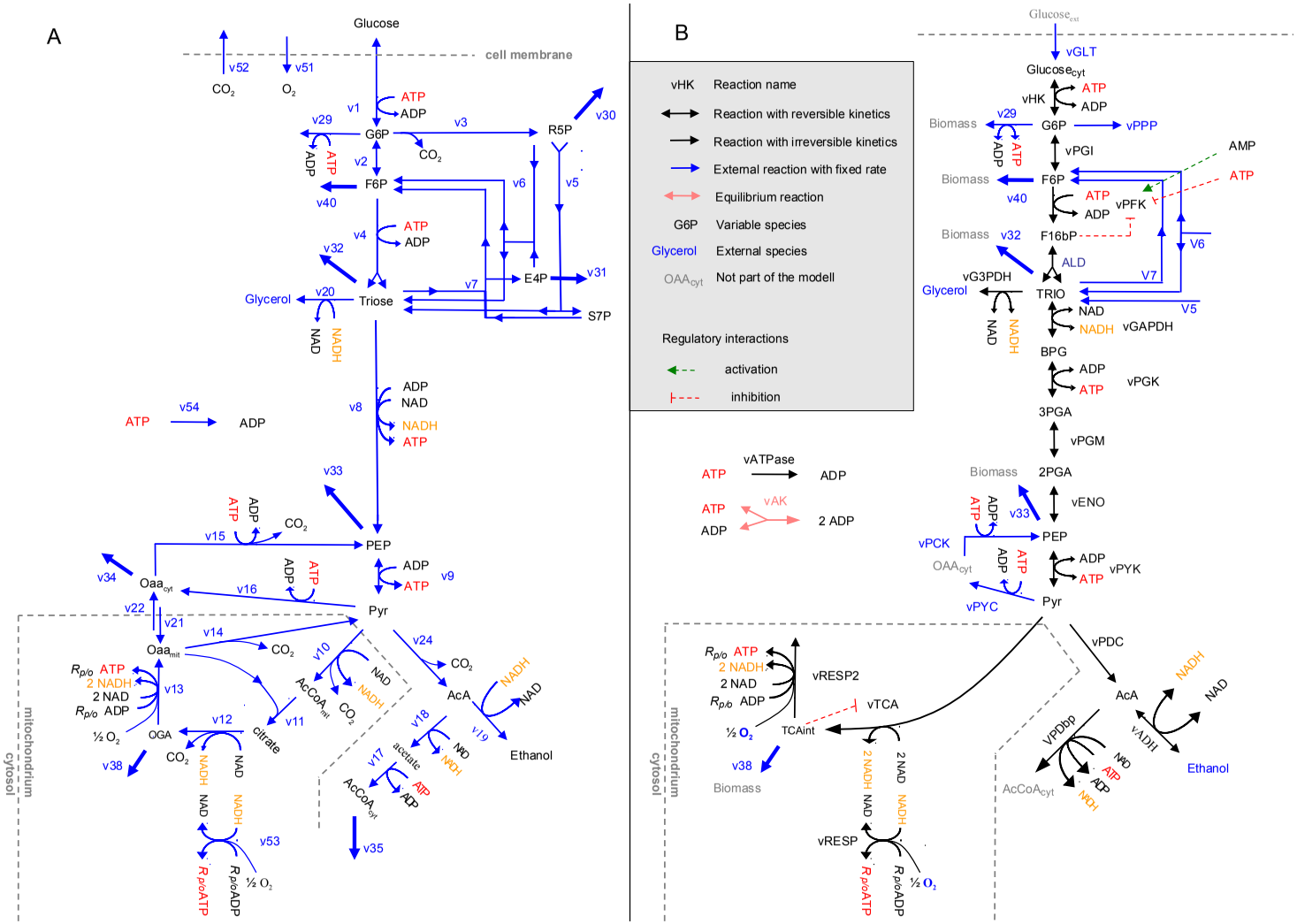

Figure 4.4: The two models of the central carbon metabolism of *S. cerevicieaes* with consistent stoichiometries. *Panel A* shows the stoichiometric network used for the constrained flux balance analysis to estimate the flux distributions shown in Fig. 4.5. *Panel B* shows the scheme of the kinetic model for which the resulting flux distribution was used. Subscripts `mit` and `cyt` denote mitochondrial and cytosolic species in the stoichiometric network. $R_{P/O}$ denotes the p/o-ratio which was set to one. The models do not possess compartments.

4.2.3 Balanced flux distribution from experimental data via constrained flux balance analysis

In order to have steady state fluxes for the stoichiometric network presented above, the constrained flux balance analysis procedure performed in [47] was partially repeated on this network. To achieve this, we used the outcome of the ^{13}C measurements in the form of the following constraints taken from [47]:

$$\begin{aligned} fr_1 &= \tfrac{v5+3\cdot v6+2\cdot v7}{v5+2\cdot v4+v6} \\ fr_2 &= v15/(v15+v8) \\ fr_3 &= v21/(v21+v13) \\ fr_4 &= v16/(v16+v22) \\ fr_5 &= v14/(v14+v9) \\ fr_6 &= v14/(v14+v9) \end{aligned} \tag{4.11}$$

where fr_i, $i = 1 \ldots 6$ are constants whose values are determined by measurements with ^{13}C supplemented chemostat cultures. The values of these are not given in [47], hence we calculated them from the flux values given in Fig. 2.2, using those sets of five values which appear on the left side.

A second set of constrains is implied by the biological knowledge that some of the reactions have zero flux under certain experimental conditions (for details s. [47]). Based on this knowledge, fluxes *v13, v15, v22* and *v51* were set to zero for the 0% oxygen condition, and the flux *v14* was set to zero for the 0%, 0.5%, and the 1% conditions.

A third set of constraints is implied by the requirement of flux balance, i.e. $N \cdot V = 0$ for the stoichiometric matrix N (shown in Appendix B) and the flux vector V.

An objective function to be minimised was constructed to include data on the external fluxes directly measured in the chemostat experiments, given by the fluxes *v1, v19, v20,* and *v51.* Since measured fluxes were given a more extensive error treatment in [47], in order to ensure good consistency with the results there, we also treated the anabolic fluxes *v29,* and *v31 - v38* as external fluxes with values taken from the publication. The these fluxes were included in the following objective function to minimise quadratic difference from the measured values while the above set of constrains were kept within a small tolerance:

$$\begin{aligned} F_{\mathrm{obj}}(V) = \; & 2 \sum_i (fr_i(V) - fr_i^{\mathrm{nom}})^2 \\ & + \sum_{j\mathrm{ext}} (V_{j\mathrm{ext}} - V_{j\mathrm{ext}}^{\mathrm{nom}})^2 \end{aligned} \tag{4.12}$$

where $fr_i(V)$ denotes the result of the evaluation of the expression fr_i in Eq. 4.11 for the flux V, fr_i^{nom} denotes its nominal value (i.e. evaluated for the flux values presented in Fig. 2.2), $j$ext denotes the indices of those fluxes which we treated as external, and $V_{j\mathrm{ext}}^{\mathrm{nom}}$ their nominal value taken from Fig. 2.2. The search for fluxes distributions V with minimal $F_{\mathrm{obj}}(V)$ under adhering to the above set of constrains was performed by the Matlab function FMINCON (as in Jouhten *et al.*). The search was repeated 100 times and the flux distribution with the smallest overall objective function selected.

The result of the constrained flux balance analysis procedure using the modified stoichiometric network is shown in Fig. 4.5. While the difference between the flux distributions presented in Fig. 2.2 and our results are minor, the flux distributions presented here are, in addition to being consistent with the measurement data in [47], also feasible steady state flux distributions for our modified stoichiometric network. As discussed in the next section, this implies that they are admitted as steady state fluxes by the stoichiometry of the kinetic model to be presented.

Compared to [47], the Fluxes *v53* (respiratory chain) and *v54* (general ATPase reaction) are new, while the Flux *v13* - as part of the respiratory chain - couples the rates of change of oxygen, NADH, and ATP.

Due to the numeric nature of the procedure, the constraints are met with finite precision. Of the above set of constraints, the most critical for our purpose is the flux balance equation $N \cdot V = 0$ which is fulfilled with a precision higher than 10^{-13} mM/min.

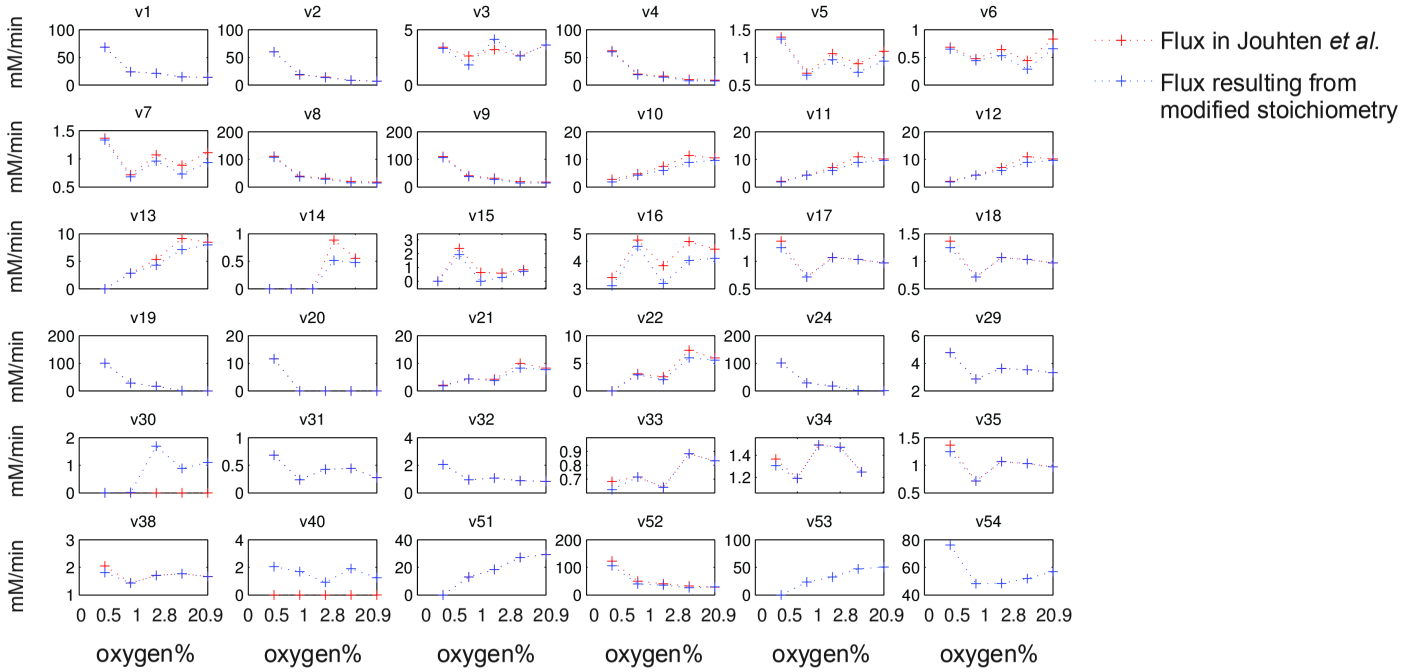

Figure 4.5: Results of constrained flux balance analysis for central carbon metabolism of *S. cerevisiae* in chemostat cultures subjected to 20.9%, 2.8%, 1.0%, 0.5% and 0.0% of oxygen in the inlet gas. Reaction names refer to Fig. 4.4. Values from the original publication [47], based on values shown in Fig. 2.2 (left row), are compared to values resulting from repeating the process using the modified stoichiometric network shown in Fig. 4.4. In comparison to the analysis in [47], reactions *v51* - *v54* were added to the original network.

4.2.4 A kinetic model for central carbon metabolism

Our aim was to construct a kinetic model which can account for both anaerobic and aerobic metabolism, and can be used together with the datasets presented in Chapter 2. As discussed in the previous sections, the modified stoichiometric network represented in Panel A in Fig. 4.4 was constructed with the primary aim to be used within the kinetic model which is presented in Panel B.

The conceptually simplest approach to construct a kinetic model based on a stoichiometric network would be to assign an appropriate kinetic expression to each reaction in the network. However, this way is seldom chosen in practice, since each kinetic expression adds to the complexity of the model, usually mirrored in the added number of parameters.

Hence, in the kinetic model, we lumped together some reactions which are distinct in the stoichiometric network. On the other hand, we wished to represent the glycolytic pathway, the core of central carbon metabolism, in more detail than in the stoichiometric network.

The above imply that the stoichiometry of the kinetic model to be presented is not the same as the stoichiometric network we used to determine the flux distribution based on data. However the two models are constructed such that:
(i) A flux distribution in the stoichiometric network can be uniquely mapped to a flux distribution in the kinetic model.
(ii) A flux distribution which is balanced stoichiometric network is balanced in the kinetic model.

Properties (i) and (ii) imply that the flux distributions shown in Fig. 4.5 may be used in the kinetic model as steady state flux distributions. This gives five different complete sets of fluxes for the kinetic model, corresponding to the five oxygenation conditions in the chemostat cell cultures.

Flux values from these flux distributions were used in two ways in the kinetic model: Some reactions were set to a constant rate of the corresponding flux value i.e. these numbers were used as parameters of the model with known *a priori* values. These reactions possess no real 'kinetic', nevertheless may influence the behaviour of the model by consuming or producing reactants. For the rest of the model's reactions, flux values were fitted, i.e. set as nominal quantities to be reproduced by the model. Since this involves a kinetic expression for the corresponding reaction, this approach was connected to one or more additional free parameters in the model. Panel B in Fig. 4.4 shows these two sets of reactions in different colours.

With the above motivations in mind, a kinetic model was constructed. The reactions of the model are listed in Table 4.2, the corresponding list of kinetic expressions is presented in Eq. B.1 in Appendix B; most of these equations are taken from [87]. The balance equations, defining the species and the stoichiometry of the model, are listed in Table 4.3. A visual representation is shown in Panel B, Fig. 4.4.

The kinetic model leans on Teusink's work [87], discussed in Section 4.1.3, which provides a list of enzyme kinetic expressions, as well as a set of corresponding parameters which were determined, as far as possible, on the basis of experimental data.

As in the Teusink model, the glycolytic intermediates GAP and DHAP are lumped together in a metabolite named Triose (Trio), and the reaction catalysed by triose phosphate isomerase (TPI) is assumed to be in equilibrium. Thus, the concentration ratio of this metabolites is fixed to the equilibrium constant $\texttt{Keq}_{\mathrm{TPI}}$, leading to the explicit formulae

$$\begin{aligned} DHAP &= \mathrm{Trio}/(1+\texttt{Keq}_{\mathrm{TPI}}) \\ GAP &= \texttt{Keq}_{\mathrm{TPI}}\ \mathrm{Trio}/(1+\texttt{Keq}_{\mathrm{TPI}}) \end{aligned} \tag{4.13}$$

Hence, time evolution of GAP and DHAP in this model is determined by the single variable $\mathrm{Trio}(t)$ and one model parameter $\texttt{Keq}_{TPI}$. Formally, these two metabolites were not explicitly part of the model equations, Eq. 4.13 was simply used in kinetic expressions requiring GAP or DHAP concentrations.

The three adenosine species (ATP, ADP, AMP) are determined by a single dynamic variable, $\mathrm{P}(t)$, and the parameters Σ_P and $\texttt{Keq}^{AK}$. This is implied by two assumptions of the model: the conservation of the sum (Σ_P) of these species, and the equilibrium for adenylate kinase reaction with (equillibrum constant $\texttt{Keq}^{AK}$). These assumptions translate into algebraic constraints (see [87]) which allow to reduce the number of independent variables and calculate concentrations of all three species from a single dynamic variable of the ODE system, $\mathrm{P}(t)$, which can be interpreted as the concentration of high energy phosphates. The resulting formulae are

$$\begin{aligned} ATP(P) &= (P - ADP)/2 \\ ADP(P) &= \frac{\Sigma_P - \sqrt{P^2\,(1-4\,\texttt{Keq}_{\mathrm{AK}}) + 2\,\Sigma_P\,P\,(4\,\texttt{Keq}_{\mathrm{AK}}-1) + \Sigma_P^2}}{(1-4\,\texttt{Keq}_{\mathrm{AK}})} \\ AMP(P) &= \Sigma_P - ATP - ADP \end{aligned} \tag{4.14}$$

Feasible range of the variable P is limited, as shown in Fig. 4.6.

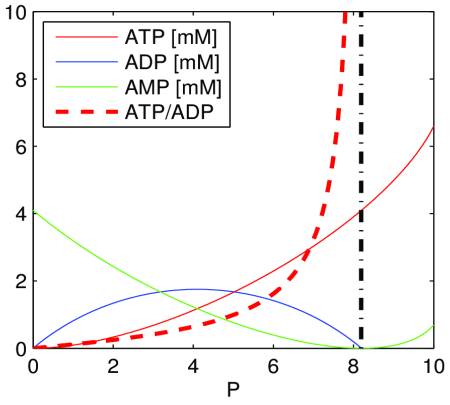

Figure 4.6: The three adenosine cofactor species and the ATP/ADP ratio as functions of the single variable P, according to Eq. 4.14. Parameters values are $\Sigma_P = 4.1$, and $\texttt{Keq}^{AK} = 0.45$, as used in [87]. Feasible range of P is limited to the range left from the black slash-dotted line, in which all concentrations have positive value. The plot also highlights the role of adenylate kinase in 'signalling' low energy states via high AMP concentrations.

The differences to the kinetic model published in [87] include the following points:

- A representation of the TCA cycle and respiration was added. In comparison to the kinetic network defined by Table 4.1, metabolic processes were further simplified or omitted.

  Reaction *vTCA* represents pyruvate transport into the mitochondria and the TCA cycle up to succinate, hence, it corresponds to the three reactions *v10, v11*, and *v12* of the stoichiometric network.

  The metabolite TCAint represents a 'generic intermediate metabolite' within the TCA cycle. It is an inhibitor of the reaction *vTCA*. In this work, the nominal concentration of TCAint was set to that of citrate. *TCAint* is also consumed by an anabolic reaction corresponding to *v38* in the stoichiometric network, the rate of which was treated as an external reaction.

  The aerobic reaction *vRESP2* corresponds to Reaction *v13* in the stoichiometric network and hence represents the reaction catalysed by the succinate dehydrogenase/complex II, as well as the corresponding proton flux through ATP synthase, resulting in ADP regeneration, and hence, creating the important stoichiometric link between these processes.

  Reaction *vRESP* represents the respiratory chain. The reaction kinetics of Reactions *vRESP* and *vRESP2* were simplified into mass-action kinetics with a single respective parameter.

  The species O_2 was introduced as an external metabolite. Since *in vivo* mitochondrial oxygen concentrations were unknown, its value was set to the value of corresponding oxygen solubility concentration given in Table 1 in [97]. It appears as a substrate in the mass-action kinetics of the Reactions *vRESP* and *vRESP2*, implying that zero O_2 concentration will halt these reactions, as it is known to be the case for the corresponding processes.

- In addition to the adenosine cofactors (ATP, ADP, AMP), the nicotinamide species NAD and NADH were introduced as dynamic variables. The sum of the these is preserved by the stoichiometry of the model, which makes the sum a conserved quantity of the model (in the sense of dynamical systems theory) and would allow to reduce the number of associated free variables to one. However, in contrast to the adenosine species, NAD and NADH concentrations were independently calculated; conservation of their sum during time evolution was observed.

- In order create a kinetic model consistent with the presented stoichiometric network, the following were added as external reactions: *v29, v32, v33, v38*, and *v40* (associated with biomass generation), as well as *v3 v5, v6, v7, v15*, and *v16* (associated with processes in the pentose phosphate pathway and gluconeogenesis).

- Within the branch towards ethanol, the model possesses a reaction *vPDbp*, representing the cytosolic pyruvate dehydrogenase bypass. (It resembles the branch towards succinate in [87], c.f. Fig. **??**) The mitochondrial part

of this pathway was concluded to be inactive during the experiments, and is not represented in the model, c.f. Section 2.4, and [47].

Lumping together a number of metabolites into the species TCAint seems an acceptable approximation in this scenario, since we did not aim to explore their individual dynamics, and the change pattern of the metabolites represented in the dataset (citrate, malate, fumarate, succinate) is similar across the five conditions, even if their concentration values differ.

The metabolite *TCAint* has two roles in this model: Because of stoichiometric constrains, at least one intermediate in the TCA cycle is necessary to enable the model to have an active TCA cycle under anaerobic conditions. As reported in the flux dataset, reaction *v38* consumes all of the TCA carbon flux under anaerobic condition when flux through the final part of the TCA cycle (starting with succinate dehydrogenase) is reported as zero.

Further, *TCAint* also plays a role in the kinetics of the model, since it acts as an inhibitor of the reaction *vTCA* (This inhibitory link is parametrised by `vTCA_KiTCAint`). This was necessary, since *vTCA*, being irreversible, is not effected by product concentration, hence, prevents the model from reaching a steady state by accumulating TCAint *ad infinitum* in a number of scenarios, if not 'explicitly' inhibited by it.

This simple regulatory loop is intended to mimic the complex regulatory mechanism of the pyruvate dehydrogenase complex which possesses a dedicated kinase-phosphatase pair, regulated by the product of the enzyme complex (c.f. Section 3.3). While this mechanism certainly allows for much higher degree of versatility than the inhibition loop in our model, one of its functions is presumably to simply prevent harmful product accumulation e.g. when oxygen provision is cut off in the environment, causing the TCA cycle reaction succinate dehydrogenase/complex II to halt.

The complete list of kinetic equations is presented in Appendix B, Eq. B.1. The the following list shows the three phenomenological kinetic expressions, used in reactions added to the model to represent aerobic energy metabolism.

$$\begin{aligned} vTCA =& \frac{\texttt{vTCA\_Vmax}}{(1+\frac{TCAint}{\texttt{vTCA\_KiTCAint}})} \cdot \frac{1}{(1+\frac{\texttt{vTCA\_KmPyr}}{Pyr})\cdot(1+\frac{\texttt{vTCA\_KmNAD}}{NAD})} \\ vRESP =& \texttt{vRESP\_K} \cdot O_2 \cdot NADH \cdot ADP \\ vRESP2 =& \texttt{vRESP2\_K} \cdot O_2 \cdot TCAint \cdot ADP \cdot NAD \\ vATPase =& \texttt{vATPase\_K} \cdot ATP \end{aligned} \tag{4.16}$$

The parameters `vTCA_Vmax`, `vRESP_K`, and `vRESP2_K` were estimated, as discussed later. As discussed below, the inhibiton parameter `vTCA_KiTCAint` could not be estimated using steady state datasets, hence its value was set to 5 mM. This value was chosen fall between concentration values for citrate (which were set as nominal values for the species TCAint) measured in the anaerobic and aerobic experiments.

Reaction name in kinetic model	Name or description of represented processes	Corresponding reactions in Table 4.1
vGLT	glucose transport	v1
vHK	hexokinase	v1
vPGI	phosphoglucoisomerase	v2
vPPP	main PPP flux	v3
v5, v6, v7	fluxes to and from PPP	v5, v6, v7
vG3PDH	glycerol-3-phosphate dehydrogenase	20
vPFK	phosphofructokinase	v4
vALD	aldolase	v4
vGAPDH	glyceraldahyde-3-P dehydrogenase	v8
vPGK	phosphoglycerate kinase	v8
vPGM	phosphoglycerate mutase	v8
vENO	enolase	v8
vPCK	phosphoenolpyruvate carboxykinase	v15
vPYK	pyruvate kinase	v9
vPYC	pyruvate carboxylase	v16
vTCA	pyruvate transport to mitochondria and TCA cycle up to OGA	v10, v11, v12
vRESP	NADH dehydrogenase, electron transport chain, ATP synthase	v53
vRESP2	TCA cycle from OGA, complex II, electron transport chain, ATP synthase	v13
vPDC	pyruvate decarboxylase	v24
vADH	alcohol dehydrogenase	v26
vPDbp	pyruvate dehydrogenase bypass	v17, v18
vATPase	ATP consuming processes	v54
v29, v40, v32, v33, v38	flux towards anabolic processes from G6P, F6P, Triose, PEP, and TCAint, resp.	v29, v40, v32, v33, v38

Table 4.2: List of reactions in the kinetic model. In addition, the reactions catalysed by adenosine kinase (AK) and triose phosphate isomerase (TPI) are implicitly represented via parametrised algebraic constraints, as they are assumed to be in equilibrium, s.text. PPP denotes the pentose phosphate pathway.

4.2.5 Representing experimental conditions

The interface between the presented kinetic model and the rest of the world is defined by so-called external[4] concentrations and external fluxes, hence, from a formal point of view, they are part of the model's assumptions. Representing steady state anaerobic and aerobic cell cultures as the "outside world" for the model was achieved by setting external quantities to appropriate values. The set of external concentrations and fluxes used in the procedures described in this chapter, is indicated in Fig. 4.4 in blue colour; differences from this list will be explicitly indicated. The values for the two modelled conditions are listed in Table B.4 in Appendix B.

Choosing the list of external quantities is part of defining the boundary of the model. Whether a quantity with known behaviour which is, at

[4]In this context, the word refers to objects being external *to the model*, rather than to the cell.

$$\frac{d}{dt}(Gluc_{\mathrm{cyt}}) = -vHK + vGLT \tag{4.15}$$

$$\frac{d}{dt}(G6P) = vHK - vPGI - v3 - v29$$

$$\frac{d}{dt}(F6P) = vPGI - vPFK + v6 + v7 - v40$$

$$\frac{d}{dt}(F16P) = vPFK - vALD$$

$$\frac{d}{dt}(Trio) = 2\ vALD - vGAPDH - vG3PDH + v5 + v6 - v7 - v32$$

$$\frac{d}{dt}(BPG) = vGAPDH - vPGK$$

$$\frac{d}{dt}(P3G) = vPGK - vPGM$$

$$\frac{d}{dt}(P2G) = vPGM - vENO$$

$$\frac{d}{dt}(PEP) = vENO - vPYK + vPCK - v33$$

$$\frac{d}{dt}(Pyr) = vPYK - vPDC - vTCA - vPYC$$

$$\frac{d}{dt}(AcA) = vPDC - vPDbp - vADH$$

$$\frac{d}{dt}(NAD) = -vGAPDH - 2\ vTCA - 2\ vRESP2 - vPDbp + vADH + vG3PDH + vRESP$$

$$\frac{d}{dt}(NADH) = \ vGAPDH + 2\ vTCA + 2\ vRESP2 + vPDbp - vADH - vG3PDH - vRESP$$

$$\frac{d}{dt}(P) = -vHK - v29 - vPFK + vPGK + vPYK - vATPase - vPCK - vPYC$$
$$- vPDbp + R_{\mathrm{p/o}}\,(vRESP + vRESP2)$$

$$\frac{d}{dt}(TCAint) = +vTCA - vRESP2 - v38$$

Table 4.3: Balance equations of the kinetic model. Additionally, the following species appear in the kinetic equations: ATP, ADP, AMP, are functions of the variable P(t), defined by Eq. 4.14. GAP, and DHAP are functions of the variable Trio(t), defined by Eq. 4.13. The species O_2 is time independent and can be treated as a parameter. Stoichiometry of the reactions *vRESP* and *vRESP2* is specific to *S. cerevicieaes*, s. Section 4.2.2. $R_{\mathrm{p/o}}$ denotes the p/o ratio which was set to one, s. text.

the same time, assumed to be influenced by the model's dynamics, is defined as external, or rather included as a variable with a defined nominal value or time course to be reproduced, is a choice of the modeller. A typical case of this decision in kinetic modelling of metabolism, is the choice whether a certain co-metabolite is internal or external to the model.

Hence, the list of external reactions was to some extent a matter of choice, especially since, as discussed before, we had detailed and consistent steady state flux distributions at hand. Which reactions we chose to be external was influenced by factors such as model complexity - e.g. including more reactions involving cofactors resulted in more difficulties in model handling - and our focus on the effect of oxygen provision in the main energy generating pathways.

For the chemostat experiments in [97], concentrations of external metabolites in the chemostat culture were given, with the notable exception of glucose, c.f. Chapter 2. The metabolites measured in the fluid phase were ethanol and glycerol. Concentration of the gaseous metabolites could not be directly measured, however, oxygen solubility concentration was given which we used as external concentration. We did not distinguish between species inside and outside the cell.

As external flux values, we used - reaping the fruit of our work - the flux distributions resulting from the constrained flux balance analysis described in Section 4.2.3. At this point, the stoichiometric consistency of the kinetic model with the stoichiometric network used in that procedure is crucial, and the main reason for carrying out the constrained flux balance analysis.

In marked difference to typical studies centred around kinetic modelling of yeast central carbon metabolism ([87], [71], [43]), we chose to make glucose transport an external reaction ($vGLT$). The motivation for this was that neither external, nor internal glucose concentration, quantities which would crucially influence the model's behaviour, were measured, and measurements from comparable aerobic chemostat studies are not known to the author at this point. As discussed in Chapter 2, data from batch cultures could not be taken as valid for this quantities. Furthermore, glucose transport in *S. cerevisieae* is a highly complex process, involving several transporter proteins of the HXT family with a variety of kinetic properties. Even though studies exist ([70], [72]), we felt that the experimental data under consideration (steady state flux and expression data, no substrate or product data) are not sufficient for modelling this process, and would divert away the focus of the work at this point.

4.2.6 Searching the parameter space to reproduce data

General considerations

In the following, we will test the model's capability to reproduce steady state values for fluxes and concentrations - the two dynamically changing quantities - as measured in the anaerobic and fully aerobic continuous cultures (0% and 20.9% oxygen in the inlet gas, respectively).

Regarding reproduction of experimental data, so far we have been mainly concerned with stoichiometric constraints: our approach guaranteed that the presented flux distributions are within the space of balanced fluxes with respect to the stoichiometry of the model. However, this is only a necessary requirement

to reproduce these flux distributions. Regarding concentrations of intermediate metabolite species, these are not determined by the above statements about the stoichiometric structure of the kinetic model, rather, depend on the kinetic rate expressions used. For example, flux through a linear pathway (or branch) in a kinetic model, containing a Michaelis-Menten-type kinetic expression, cannot exceed the $V_{\max}$ value used there, even if stoichiometric constraints would permit it.

Kinetic laws in this model possess a high number of parameters. With regard to finding sets of parameter values of the model which reproduce flux and concentration values associated with experimental measurements, the following aspects are to be be discussed:

(i) Suppose that the stoichiometry of the model allows for the desired objective flux distribution, do indeed sets of kinetic parameter values exists such that the model produces this distribution ?

(ii) If so, to what extent can this be done with the concentration of intermediate metabolites being in an appropriate range?

(iii) If (i) and (ii) are achieved, to what extent are the corresponding parameter sets unique, and - connected to this question - can they be usefully interpreted as 'predictions'?

We tackled (i) and (ii) at the same time using a parameter estimation technique with an objective function in which the 'distance' of concentrations, fluxes and other quantities from their desired values or ranges were taken into account. The parameter search was limited by the high number of parameters in the model.

Nominal values and nominal ranges to be reproduced

Our aim was to reduce ambiguity in the parameter set of the model as much as possible by providing a high number of objectives to be reproduced by the model. For the $V_{\max}$ of the reaction kinetic expressions to be defined, we aimed for defining a full set of nominal metabolite concentrations based on experimental data.

Quantities to be reproduced (fitted) by the model were assigned a nominal value as well as a nominal range within which the concentration was allowed to vary. The aim of this approach is to represent measured quantities with uncertainties in the parameter estimation process. During this procedure described in Section 4.2.6 below, parameter sets were assigned a 'goodness of fit' value, constructed by assigning penalties for each fitted quantity based on how well it is reproduced by the model. Parameter sets resulting in simulations which exactly reproduced a nominal value were assigned no penalty for corresponding quantity, however even if this was not the case, but the corresponding variable was still within the nominal range, the associated penalty for the corresponding parameter set was low. We used this scheme as implemented in the software tool SBTOOLBOX2 [77].

The nominal values and ranges for various measured quantities are indicated in Fig. 4.8; details are discussed below.

Intermediate metabolites contained in the metabolite dataset discussed in Chapter 2, nominal concentrations were simply defined by the mean value and standard deviation[5].

The metabolite dataset used so far in this work does not contain a full list of metabolites. The intermediate metabolite species Trio (sum of GAP and DHAP), G3P, G2P, BGP of the presented kinetic model are not directly measured. For the species TCAint (representing a number of internal metabolites in the TCA cycle) the nominal value for citrate was used.

In order to set nominal concentration ranges for the missing metabolites literature data were used. This proved difficult, since only a few published studies including the above species were found. In addition, studies using anaerobic cultures, especially anaerobic chemostat cultures are less numerous. Unsurprisingly, studies using cultures of a yeast cell line closely related to the one used in [97], targeting the above metabolites, were not found.

Publications using chemostat cultures grown in microaerobic conditions were not found - this fact was a motivation for the study presented in Chapter 2 - which was one of the reasons why we only aimed to reproduce the anaerobic and the anaerobic data at this point.

The size of the nominal concentration range for a species reflects our opinion about the uncertainty of that value, taking into account how similar the experimental conditions were to that of the experiments discussed in Chapter 2.

Hence data taken or inferred from [90, Theobald *et al.*] for the aerobic experiment were assigned larger nominal range than typical standard errors of concentration data in [97], since the cell line and presumably the level of oxygenation of the culture were different. Even larger nominal ranges were assigned to values taken or inferred from [87, Teusink *et al.*] for the anaerobic experiment, since this study was based on shake flask cell cultures as opposed to our chemostat cultures. However, in some cases information from more than one of these studies had to be used to estimate a concentration value for a metabolite. A large nominal range was assigned to the concentration of the species BPG, since this is not measurable due to its low *in vivo* value. We used the value indirectly inferred by [87].

Our resulting choice of nominal values and associated nominal ranges is shown in Table B.3 in Appendix B.

Fluxes were assigned somewhat smaller nominal ranges - of about one fifth of their respective value - than the typical range for the concentration data taken from [97], since the ^{13}C constrained MFA procedure used in [47] proved more robust than the concentration measurements. Moreover, we assigned higher importance to reproducing flux distribution than metabolite concentrations. Assigning smaller ranges to the flux data allowed to encode this preference in the objective function of the parameter search method.

Each flux was assigned the corresponding nominal value (f_{nom}) from Section 4.2.3 and the nominal range $[f_{\text{nom}} \cdot \frac{5}{6},\ f_{\text{nom}} \cdot \frac{6}{5}]$.

Co-metabolite concentrations were assigned nominal ranges via the following procedure, aiming at defining a range of co-metabolite ratios to be produced by the model, rather than specific values. The momentarily values of

[5] In contrast to the standard error of the mean used in the original publication [97]

the three adenosine and the two nicotinamide co-metabolites are determined or may be reduced to a respective single variable in the presented kinetic model. Hence, a further algebraic constraint fully defines their values.

Instead of aiming to reproduce concentrations, this enabled us to to assign nominal ranges to the concentration ratios $R_1 = ATP/ADP$ and $R_2 = NAD/NADH$, since we felt that this quantities constitute a more important characteristic of the model, and hence, should control the parameter search in a direct manner.

Encoding this in the objective function required to construct auxiliary model variables $F_i(R_i)$ with the property that they allow variation of the respective ratios within a certain range without much variation, but change quickly outside these ranges. We then used these variables as quantities to be fitted.

Ratios of non-negative quantities are more appropriately handled in logarithmic space, hence we required that deviation of the ratio R_i, $(i = 1, 2)$ from the targeted nominal value $R_{i\,nom}$ fulfils the requirement

$$F_i(x \cdot R_{i\,nom}) = -F_i(\frac{1}{x} \cdot R_{i\,nom}),\ i = 1, 2 \tag{4.17}$$

where F_i denotes the relevant part in the objective function with the property $F_i(R_{nom}) = 0$ which follows from the above equation for $x = 1$.

The above motivation lead to the following construction for the F_i:

$$F_i(R_i) := \left(\frac{log(R_i/R_{i\,nom})}{log(C_{i\,nr})}\right)^{n_i} \tag{4.18}$$

where $C_{i\,nr} > 1$ and $n_i \in N$ denote constants defining the nominal range and the steepness of the function F_i outside this range, as discussed below. As required, this function gives zero for $R_i = R_{i\,nom}$. The role of the other constants becomes clear by noticing that

$$\begin{aligned} F_i(C_{i\,nr} \cdot R_{i\,nom}) &= 1 \\ F_i(R_{i\,nom}/C_{i\,nr}) &= -1 \\ &\text{and} \\ |F_i(R_i)| &\leq 1 \quad \text{for } R_i \in [R_{i\,nom}/C_{i\,nr},\quad C_{i\,nr} \cdot R_{i\,nom}] \\ |F_i(R_i)| &> 1 \qquad\qquad\qquad \text{otherwise} \end{aligned} \tag{4.19}$$

The absolute value of F_i is smaller than unity for concentration ratios R_i within the range defined by $C_{i\,nr}$, and larger outside. By choosing a larger exponent n_i, this behaviour can be amplified such that F_i will have very small absolute values for most R_i within the defined range and exhibits a steep rise as R_i crosses the border to the outside. If n_i is chosen to be an odd number, $R_i < R_{i\,nom}$ will give negative values, such that F_i becomes a monotonous function.

The behaviour of the variables F_i as functions of the metabolite ratios is shown in Fig. 4.7.

It follows from the above that the variables F_i map co-metabolite ratios within their respective nominal ranges into the range [-1, 1]. Hence, we assigned zero as nominal value and [-1, 1] as nominal range to the variables F_i (i=1,2) for the parameter estimation.

The means: estimating parameters

The kinetic expressions in the model contain approximately ninety parameters in total, most of which were adopted from [87]. In view of the size of our dataset, and due to the fact that the corresponding experiments corresponded to steady states on the timescale of metabolic reaction kinetics, we concluded that estimating a considerable part of these parameters is impossible within the scope of this work. However, as mentioned above, Teusink's model already contains a set of parameters which is - to a high degree - based on experimental data.

Therefore, we tackled the parameter estimation problem by asking, which part of the cell's physiology is likely to show clear difference between the cell cultures used in [87] and the cultures used in [97],[47], [69] to generate our datasets. It seemed that the most natural candidates are the capacities of metabolic enzymes. This reflects the assumption that the abundance, and hence, capacity of many metabolic enzymes will generally differ between different strains or between cultures grown under different conditions, while other parameters encode properties of the enzymes itself, and as such, were assumed to be (approximately) constant.

The parameters representing the *in vivo* enzyme capacities are the $V_{\max}$ values in the kinetic expressions used to describe reaction rates. Therefore, regarding kinetic expressions taken from [87], our approach was to adjust $V_{\max}$ values with the aim to reproduce experimental data, while leaving all other parameters in these kinetics unchanged. A second set of parameters to be estimated were those contained in the phenomenological kinetic expressions for the reactions *vTCA*, *vRESP*, *vRESP2*, and *ATPase* since, due to their highly simplifying nature, no experimental literature values could be used. In addition we estimated parameters and initial values connected to the adenosine and nicotinamide cofactor species. The aim was to estimate the above sets of parameters for aerobic and anaerobic conditions individually and compare the results.

Due to the high number of parameters (>15), search range for individual

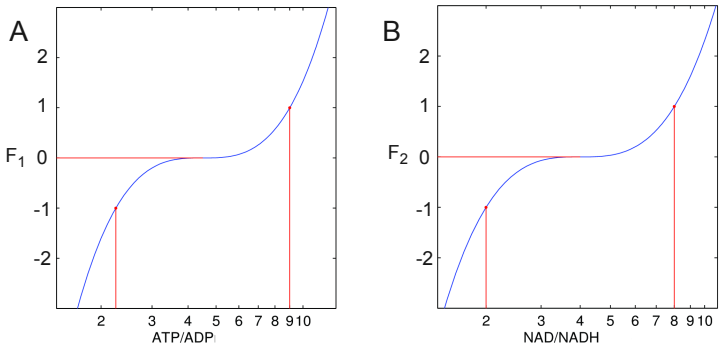

Figure 4.7: Graph of the auxiliary variables F_1 and F_2 as functions of the corresponding co-metabolite ratios as defined by Eq. 4.18. These variables were constructed in order to be used as fitted quantities during parameter estimation. They map the predefined range of the respective co-metabolite ratio onto the interval [-1 1] and exhibit little variation within it, but rise or fall steeply as the argument approaches the edge.
Panel A shows a graph of F_1 with nominal value $R_{1\,nom} = 5$, range $[2\frac{1}{4}, 9]$, and $n_1 = 3$.
Panel B shows a graph of F_2 with nominal value $R_{2\,nom} = 4$, range $[2, 8]$, and $n_1 = 3$.

parameters had to be chosen relatively small (typical fold changes from minimal to maximal values were 1.5 -3 fold). Choosing too large search ranges quickly resulted in subset of parameter space[6] too large for our computational capabilities. The consequence was that a parameter value set with good data fit was not found, even if previous estimations (with smaller search range) showed that such parameter sets were included in the search range. Typical amount of calculation included probing $10^5 - 10^7$ sets of parameter values corresponding to a few hours of running the parameter search algorithm.

Truly global parameter search was not possible in this scenario. However, this was not necessary for all parameters, since the order of magnitude for the values of the $V_{\max}$ were assumed to the same as in [87]. We followed two strategies to achieve parameter search as global as possible with the given limitations.

First, the searched subset of parameter space was adjusted iteratively with the aim to keep it as large as possible while still small enough to be searched with the computational resources available, and, at the same time, include all likely best fit parameter value sets. The procedure was as follows: during a high number of estimation attempts, search range for each parameter was adjusted by hand such that best fit values did not appear at its edge. Thus, if an estimation procedure resulted in a best fit value of a parameter too near (within approximately 3%) to one of the limits of its search range, the search range was slightly extended to this direction. If repeating this procedure a number of times resulted in a subset of parameter space of such large volume that it lead to a decreased ability of the algorithm used to find good fits, parameter search ranges were 'clipped' at the end which was further away from typical best fit values of the corresponding parameter. Since parameters turned out to be reasonably well defined by the estimation procedure, this did lead not a decrease of the ability of finding parameter sets with good fits. It should be noted that the above iterative procedure needs to be repeated if a component of the model with sufficient global impact is changed.
The high number of parameters together with the fact that largely different $V_{\max}$ values were necessary for reproducing anaerobic and aerobic conditions prevented the usage of the same search range for all parameters for both anaerobic and aerobic conditions. Hence, search ranges for most parameters representing $V_{\max}$ of an enzyme kinetics expression were individually adjusted for both conditions.

As the second method to reduce the dimensionality of the search space, not all branches of the model were fitted simultaneously. As an intermediate stage, rates of the reactions *vG3PDH, vPDC, vADH*, and *vPDbp* (those associated with anaerobic metabolism) were fixed to their nominal flux values, leaving, essentially, a linear pathway, corresponding to glycolysis and the reactions associated with mitochondrial metabolism. The bulk of parameter search simulations were performed in this restricted model.

It should be noted, that in the presented setting - with a detailed stoichiomet-

[6] The vastness of a high-dimensional space may be demonstrated by calculating that doubling the individual search ranges in a 10-parameter search results in increasing the volume of the associated hypercube by $2^{10} = 1024$ fold. This gives the increase of computational time when using a simple Monte-Carlo algorithm, while for more specialised algorithms the increase in computational time is expected to be less.

ric network at hand, defining nominal flux distributions for the stoichiometric model - this approach of 'stepwise estimation' may be applied to extend the kinetic model, provided, the steady state data can be fitted by the model. This way, replacing more and more fixed-rate-reactions by kinetic equations is possible, such that only parameters corresponding to a few reactions need to be adjusted at a time to fit steady state data.

This marked difference in the parameter values in $V_{\max}$ eased the interpretation of the results of the parameter search. To grasp the uncertainty in the parameter values we repeated the parameter search a high number of times and analysed those results which had a satisfactory goodness of fit (the objective function evaluated below 0.1). This resulted in the realisation that the $V_{\max}$ are mostly well defined in comparison to the difference between the two conditions, while the other parameters - mostly corresponding to the phenomenological reactions - did not exhibit reliably different values between the two conditions. This was to be expected, given that this model is overparametrised for the amount of fitted quantities. These intermediate results are shown in Figure B.1 in Appendix B.

From the 22 prefitted results, we selected a dataset for the anaerobic and the aerobic condition respectively. Simply selecting by the lowest value of the objective function was prevented by low ATP/ADP ratios in some parameter sets, hence we took into account the value of this quantity as well.

On the two selected models we performed a final estimation procedure, during which rates of the reactions *vG3PDH, vPDC, vADH*, and *vPDbp* were defined by the dynamics of their assigned kinetic expressions, estimating the corresponding $V_{\max}$ values. Other parameters were searched in a restricted search range of only $\pm$ 10% of their former optimal value.

In order to select for parameter sets resulting in higher steady state ATP/ADP ratio, this second estimation procedure was performed with corrected nominal values and ranges for the cofactor ratios. The settings for the auxiliary variable F_1 and F_2 (c.f. Eq. 4.18) were:
$R_{1\,nom} = 5$, $C_{1\,nr} = 2$, and $n_1 = 5$, corresponding to the nominal ATP/ADP range $2\frac{1}{2} - 10$.
$R_{2\,nom} = 7$, $C_{2\,nr} = 3$, and $n_2 = 3$, corresponding to the nominal NAD/NADH range $2\frac{1}{3} - 21$.

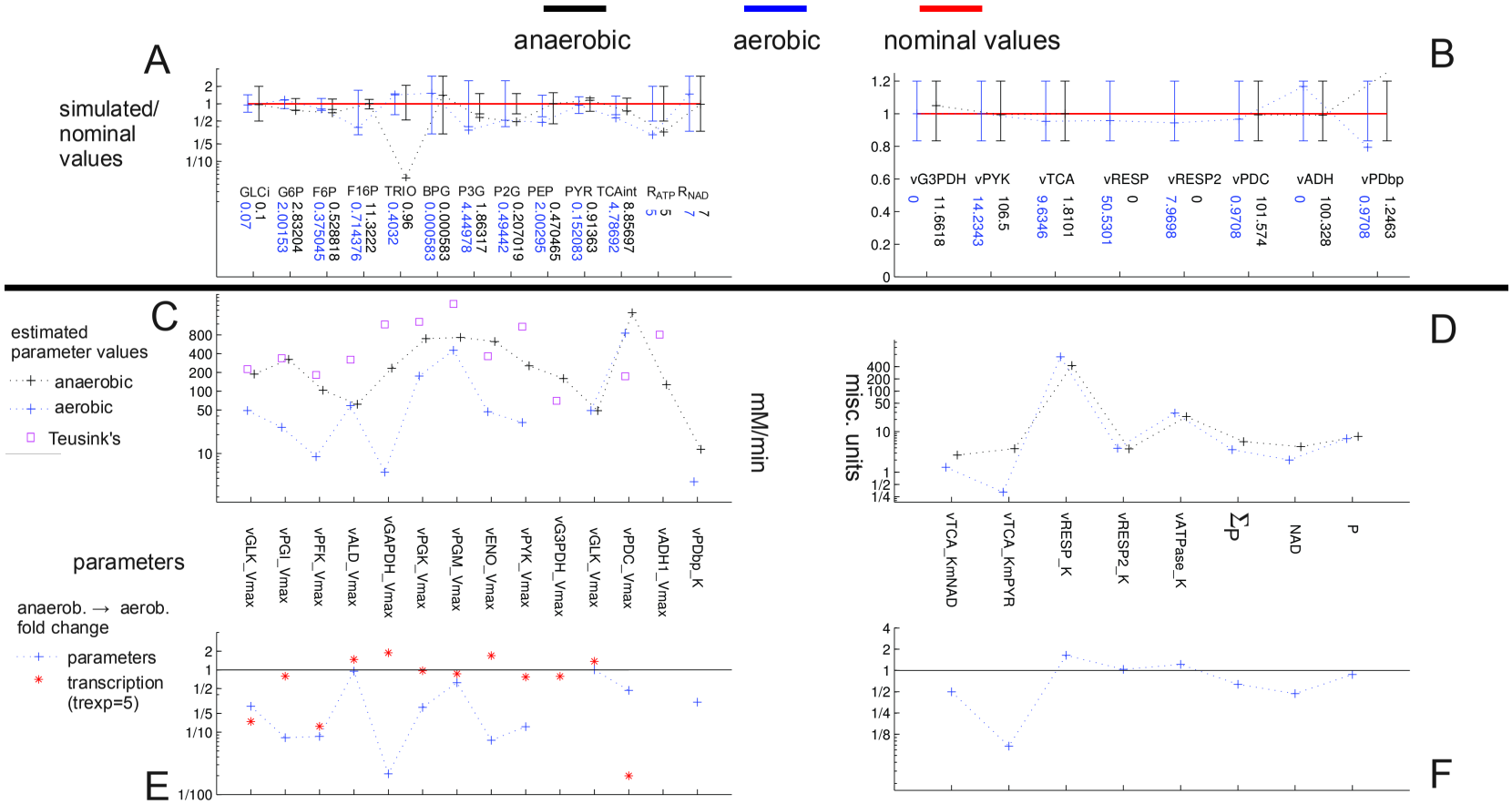

Figure 4.8: Result of fitting the kinetic model using experimental datasets from anaerobic (0% oxygen, black) and aerobic (20.9 % oxygen, blue) chemostat yeast cultures. Shown are simulation results for the two corresponding parameter sets.
Panel A, B: Fitted quantities: Simulated concentration (Panel A, logarithmic scale) and flux values (Panel B) plotted as multiples of corresponding nominal values (appearing as unity, red line). Names and absolute nominal values (colour coded) are indicated. R_{ATP} and R_{NAD} denote ATP/ADP and NAD/NADH ratio, respectively. Uncertainties in nominal values are denoted by error bars. Error bars and simulated values are colour coded to indicate oxygenation conditions.
Panels C, D, E, F: Estimated parameters. Panels C and D show absolute parameter values plotted on a logarithmic scale. For glycolytic V_{max} (Panel B), the corresponding values used in [87, Teusink] (based on anaerobic batch cultures) are shown. *vG3PDH_Vmax* and *vADH_Vmax* were set to zero for modelling the aerobic experiment, hence, these values are not plotted.
Panels D and E show the aerobic parameter values plotted on a logarithmic scale normalised to the anaerobic parameter set. Red stars denote fold change calculated from transcription data under assumption of concerted regulation with the transcription-amplification parameter `trexp` =5.

4.3 Results and discussion

4.3.1 Regulation between anaerobic and aerobic conditions

The anaerobic and the aerobic parameter sets for the model which resulted from the above estimation procedure exhibited improved fit to the corresponding steady state datasets. The parameter sets and the corresponding steady states of the model are shown in Fig. 4.8.

As seen in Panle C, the estimation procedure resulted for most $V_{\max}$ in clearly differentiable values between the two steady states. Since these parameters represent experimentally measurable quantities, we can interpret them as predictions of the model. Numerical values of $V_{\max}$ and fold changes between the two states are presented in Table 4.4.

Their estimated values may be interpreted to be consistent with all model assumptions and the experimental information used: the metabolite and flux datasets presented in Chapter 2, as well as biochemical knowledge contained in the kinetic equations adopted from [87]. It is notable that such a set of $V_{\max}$ values exist, and that they differ sufficiently between the two steady states.

One measure of consistency is the quality of data fit. While this may be called satisfactory for the case of the two steady states presented, the concentration for the species Trio (representing the sum of the metabolites GAP and DHAP) exhibits a large discrepancy from the nominal value for the anaerobic dataset. Possible causes for this discrepancy includes a change in the kinetic properties of one of the enzymes between anaerobic an aerobic conditions which cannot be represented by changing $V_{\max}$.

This implies a lower reliability for the predicted anaerobic $V_{\max}$ values for reactions consuming or producing this metabolite *vG3PDH*, *vALD*, *vGAPDH*. However, since concentrations adjacent to Trio are close to their nominal values, the consistency for the rest of the model is assumed to be independent from this.

The largest differences in enzyme capacities are predicted for the upper glycolytic reactions associated with *vGLK*, *vPGI*, *vPFK* and *vGAPDH* in the model. On the other hand the reactions *vPGM, vGLK*, and *vALD* are not required to change in capacity, i.e. flux change through these reactions is achieved by change in metabolite concentrations. No valid fold change value was indicated for *vG3PDH* and *vADH*, since $V_{\max}$ values for these reactions were set to practically zero to fit the zero flux in the aerobic dataset.

It seems advisable to treat the predicted absolute $V_{\max}$ values with care and focus on fold changes, instead, as predicted quantities. Enzyme kinetic equations always represent simplified idealisations of complex processes, and while they may reflect the general form of the real kinetics, specific values for maximal velocities are less robust than their fold changes between conditions. In addition, a similar statement holds for the corresponding experimental data: while *in vitro* measurements have been standard procedure for decades, the resulting enzyme capacity values seem to be far from the corresponding *in vitro* values [95]. While the latter are not directly measurable presently (c.f. [25]), fold changes between conditions measured *in vitro* are thought to reflect *in vivo* values with reasonable accuracy [9].

model reaction	Vmax anaerobic	Vmax aerobic	Vmax fold change	transcription fold change	trexp
vHK	188	48.9	0.26	0.68	3.53
vPGI	324	26.3	0.08	0.95	54.04
vPFK	104	8.88	0.09	0.66	5.89
vALD	61.7	58.7	0.95	1.08	-0.65
vGAPDH	234	5.03	0.02	1.13	-30.42
vPGK	699	175	0.25	0.99	230.45
vPGM	731	455	0.62	0.97	16.25
vENO	630	46.8	0.07	1.11	-24.84
vPYK	256	31.4	0.12	0.95	39.53
vG3PDH	160	0	0	0.95	
vPDC	1823	854	0.47	1.07	-12.04
vADH	128	0.054	0.0004	0.46	9.9183

Table 4.4: Fold changes of maximal velocities (`Vmax`, [mM/min]) of reactions of the kinetic model, compared to fold changes in transcript levels coding for the corresponding enzymes between anaerobic and aerobic steady states. From these, the transcription amplification parameter `trexp` is calculated using Eq. 4.20. A negative `trexp` value indicates increase in one quantity and decrease in the other.

Transcript level changes were compared with $V_{\max}$ fold changes as follows. For each reaction, we calculated an associated transcription fold change value from the transcription data of the appropriate genes. (The details are shown in Eq. B.2 in Appendix B). Rather than directly comparing the predicted fold change values with that of transcription data, we used the approach developed in Section 3.2.2. This employs the assumption of *concerted regulation* of the various levels of hierarchic regulation for an enzyme's capacity, implying the following one-parameter formula (c.f. Eq. 3.18), relating fold change in $V_{\max}$ with that in the transcript level (denoted by `mRNA`) between the two steady states, denoted by subscribts `anaer` and `aer`:

$$\frac{V^{\max}_{\mathtt{aer}}}{V^{\max}_{\mathtt{anaer}}} = \frac{\mathrm{PROT}_{\mathtt{aer}}}{\mathrm{PROT}_{\mathtt{anaer}}} = \left(\frac{\mathrm{mRNA}_{\mathtt{aer}}}{\mathrm{mRNA}_{\mathtt{anaer}}}\right)^{\mathtt{trexp}} \tag{4.20}$$

where, under the above assumption, the parameter `trexp` is expected to roughly reflect the number of regulatory levels in hierarchic regulation of $V_{\max}$.

From predicted fold changes $V_{\max}$ and transcription data, we calculated what value the parameter `trexp` would have, if the assumption of concerted regulation was correct. The results are summarised in Table 4.4.

Since we assume the number of regulatory levels from transcription initiation to protein to be less than 10, we expect `trexp` values below this number. Only the enzymes associated with the model reactions *vHK*, *vPFK*, *vALD*, and *vADH* have `trexp` values below 10, indicating that their associated transcript fold changes are roughly consistent with the corresponding $V_{\max}$ change required for the model to fit the experimental data, under the above assumption. As remarked above, the prediction for *vALD* schould be treated with care.

Consequently, for all other reactions, the model predictions for $V_{\max}$ fold changes are not consistent with the transcription data under the assumption of concerted regulation. It is not necessarily clear where the discrepancy comes

from: false predictions for $V_{\max}$ fold changes or falseness of the assumption of concerted regulation, or both. However, transcription data are known to be unreliable predictors for enzyme capacity change which may be rephrased as the falseness of the assumption of concerted regulation.

Hence we conclude that transcript levels are not the adequate measured quantity to quantitatively compare model predictions on fold changes of enzyme capacities in this experiment.

model reaction	$J_{\texttt{anaer}}$	$g_{\texttt{anaer}}$	Vmax (anaer)	$J_{\texttt{aer}}$	$g_{\texttt{aer}}$	Vmax (aer)	R_M	R_H	ρ_m	ρ_h
vGLK	68.3	0.363	187	13.9	0.283	48.9	0.173	0.83	0.155	0.844
vPGI	60.2	0.186	323	6.9	0.262	26.3	-0.252	1.25	-0.160	1.16
vPFK	60.1	0.579	103	7.3	0.818	8.88	-0.254	1.25	-0.163	1.16
vALD	60.1	0.973	61.7	7.2	0.123	58.6	0.968	0.032	0.976	0.024
vGAPDH	106	0.454	234	14.3	2.852	5.03	-3.116	4.11	-0.916	1.92
vPGK	106	0.152	699	14.3	0.081	175	0.333	0.666	0.309	0.691
vPGM	106	0.145	731	14.3	0.031	455	0.734	0.265	0.763	0.236
vENO	106	0.169	629	14.3	0.306	46.8	-0.504	1.50	-0.296	1.30
vPYK	105	0.413	256	14.2	0.453	31.4	-0.063	1.06	-0.046	1.05
vPDC	100	0.055	1823	0.938	0.001	854	0.726	0.273	0.838	0.162
vADH	99.4	0.778	128	0.167	3.09	0.05	-1.486	2.48	-0.216	1.22

Table 4.5: Hierarchic analysis of the flux regulation between simulated anaerobic (`anaer`) and aerobic (`aer`) steady states. J denotes steady state flux, `Vmax` denotes the corresponding parameter in the kinetic expression used in the model. The quantities g, R_M, R_H, ρ_m, and ρ_h are defined in Eqs. 3.3 and 3.8 in Chapter 3.

Hierarchic regulation analysis was performed in order to elucidate the respective contributions to flux changes from the hierarchic and the metabolic regulatory levels between the presented steady states. This amounted to applying analytical methods discussed in Section 3.2.1. The regulation coefficients quantifying contributions from the hierarchic and the metabolic levels were calculated according to both formalisms introduced in this work. The coefficients according to Eqs. 3.3 (taken from [53]) are denoted by ρ_h, and ρ_m, while the alternative coefficients defined by Eq. 3.8, introduced in this work, are denoted by R_M and R_H. In order to ease their interpretation, their values for a number of scenarios were presented in Table 3.1. Calculated values for the two simulated steady states under consideration are shown in Table 4.5.

Regarding the comparison of the two steady states presented this chapter, the model reactions *vGLK, vALD, vPGK, vPGM,* and *vPDC* possess both hierarchic and metabolic coefficients greater than zero. As discussed in Section 3.4, this can be interpreted as the statement that these reactions are cooperatively regulated by the two regulatory levels.

For reactions *vPGI, vPFK, vGAPDH, vENO, vPYK*, and *vADH*, the calculated hierarchic coefficients have positive values, while metabolic coefficients are negative, corresponding to the interpretation that these reactions are antagonistically regulated, with hierarchic regulation prevailing against metabolic regulation in the flux change. In other words, the flux is changed by up- or downregulating $V_{\max}$, against the effects of chemical potential and allosteric regulation (if present). This can be interpreted such that these enzymes are the contact points at which the hierarchic regulatory level is most crucially involved in changing the flux distribution.

Apart from [53], the Westerhoff group published calculations of hierarchic coefficients based on experimental studies [73] (s. also supplement to [9]). The results presented in Table 4.5 can be best compared to Table 2 in [73]. The latter is based on an experimental study comparing three cell states corresponding to different experimental conditions. Among the reactions included in both studies, only one flux, that through ADH, is antagonistically regulated in all three of the pair-wise comparisons of steady states presented in these studies.

This is in line with our simulated results, since *vADH*'s hierarchic coefficients have the second-highest values in 4.5, showing that regulation of this reaction is not only to a high extent hierarchical, but prevails against a high extent of metabolic regulation in the model.

4.3.2 Exploring model dynamics - an *in silico* perturbation experiment

The presented kinetic model was constructed purely by taking its steady states into account. Naturally, kinetic models are expected reproduce or at least elucidate the dynamics of the system they represent. A thorough analysis of the dynamic behaviour of the presented kinetic model as well as further parameter search with the aim of reproducing known behaviour of the system is not within the scope of this present work. However, in an exploratory manner, we aimed to elucidate on the question, to what extent the presented kinetic model - without further modifications - produces expected qualitative behaviour of central carbon metabolism.

Since dynamic variables of the model are metabolic concentrations, and enzyme capacities are represented by parameters, dynamics of the model corresponds to the regulatory regime on the metabolic level. Experimentally, as discussed in the introduction of this chapter, this dynamics may be best studied within the first hundred seconds after introducing a sudden a perturbation, as performed in [90] for glycolysis (see also [91], [72], [93]). In the given context, we chose to perform an *in silico* perturbation experiment mimicking oxygen level change. The qualitative behaviour associated with such change in external conditions is to observe a redistribution of fluxes between the pathways associated with anaerobiosis (ethanol and glycerol producing branches) and the respiratory chain.

Due to the fixed enzyme capacities in the model, the aerobic parameter set presented in 4.8 is not capable to represent significant ethanol or glycerol production, since the $V_{\max}$ of the corresponding reactions (vADH and vG3PDH) are near zero.

Hence we presented the anaerobic model with a sudden introduction of $O_2^{\text{pert}} = 0.25$ mM of oxygen (represented by the species O_2), the same concentration we used to characterise the aerobic condition. All other parameters were left unchanged. To a certain extent, this scenario mimics the perturbation experiment leading to the observation of what is historically known as Pasteur effect, the drop of glucose consumption by an anaerobic yeast culture upon aeration. However, since the reaction representing glucose influx is fixed in the model, this cannot be reproduced in a direct manner. Nevertheless, we can pose the question, to what extent the other major effect, redirection of flux towards respiration, is reproduced by the model.

In order to avoid numeric complications by abrupt changes, we used a smooth function of time given by

$$O_2(t) = O_2^{\text{pert}} \frac{x(t)}{x(t)+1} \quad \text{where} \quad x(t) := exp(t-10) \tag{4.21}$$

The timecourse of $O_2(t)$ defined by the above equation and the resulting change of other species and independent fluxes is shown in Fig. 4.9.

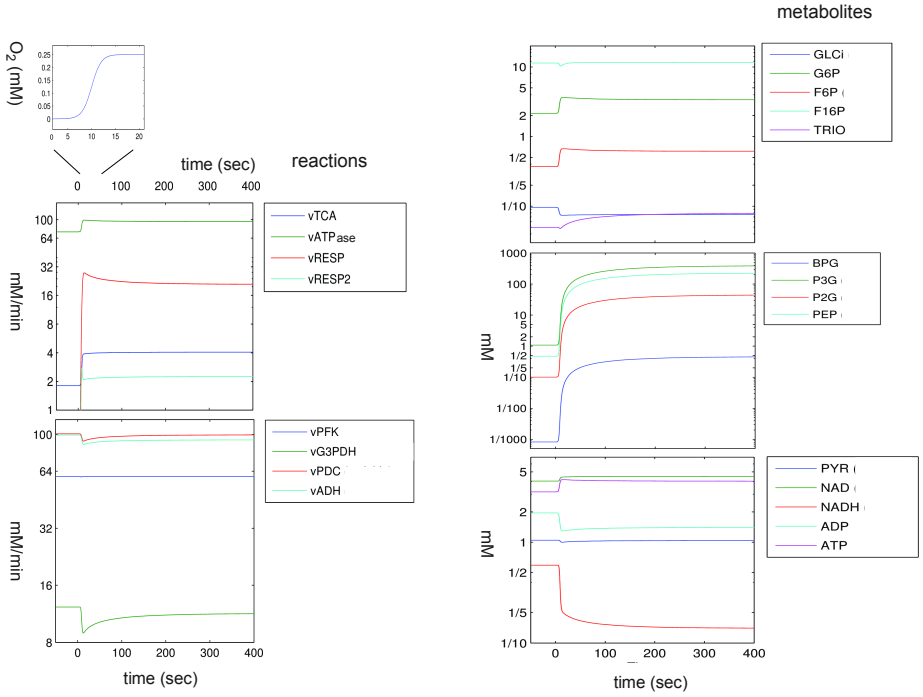

Figure 4.9: Change of concentrations and reaction rates in the anaerobic steady state upon changing the oxygen concentration to its level in the aerobic culture. Reaction rates of *vRESP* and *vRESP2* are zero before the oxygen is introduced, hence not visible in the logarithmic plot.

Flux redistribution after the perturbation showed unsurprising sudden activity of the reactions *RESP* and *RESP2* which were inactive prior to the perturbation, due to lack of their substrate, O_2. This leads to a twofold increase of the flux through *vTCA*. In turn, this leads to an increase of both the ATP/ADP, as well as the NAD/NADH ratio: ATP concentration rises from 3.2 mM to 4.0 mM while ADP falls from 2.0 mM to 1.4 mM, changing the ATP/ADP ratio from 1.6 to 2.9; NADH changes from 0.6 mM to 0.14 mM, while NAD level increases from 4 mM to 4.5 mM, changing the NAD/NADH ratio from 6.8 to 32. (approximate values).

Interestingly, the reactions towards glycerol (*vG3PDH*) and ethanol (*vADH*) exhibit a decrease, even though *vG3PDH* recovers to a certain extent. This is may seem surprising at first, since concentrations of their substrates - Trio and ACA, respectively - increase, while their products are both kept at constant

levels as external quantities. Since these reactions are not regulated by further model variables, the observed flux decrease can be accounted to the concentration changes of the second substrate-product pair, NADH - NAD. We will term this effect 'metabolic downregulation', using the nomenclature introduced in Chapter 3. As a result, the model exhibits a certain flux redistribution towards respiration.

Among intermediate metabolite levels, only the TCA-cycle intermediate TCAint and the glycolytic intermediates GLCi (internal glucose) and Pyr exhibit a decrease. Decrease in TCAint is more than 10-fold and results from the activation of Reaction *vRESP2*. Since this metabolite acts as an inhibitor of the irreversible reaction *vTCA*, its decrease can be identified as the cause for the increased flux through the latter, given that the two substrates of this reaction - Pyr and NAD - decrease.

Decrease in GLCi and Pyr can both be accounted to the change in ATP - ADP levels. As metabolic regulators, they upregulate *vHK* and downregulate *vPYK*. This can be concluded as follows: GLCi is product of the external reaction *vGLT* which, kept at a fixed rate, forces the reactions *vHK* and *vPGI* to recover to the steady state flux prior to the perturbation. Hence, decrease of GLCi concentration must be accounted to the short-lived rate increase of its sole consuming reaction, *vHK* (not shown). This, in turn, must be caused by change in the level of the adenosine cofactors, given that concentration of its product, F6P, increases. Similarly, the increase of PEP and the decrease of Pyr identifies metabolic regulation of *vPYK* by the adenosine cofactors as the only possible cause for these changes.

Quantitatively, the changes in fluxes and concentration seem to be different from what we expect from real yeast metabolism. The extent of flux redistribution is moderate, with flux changes in the order of 20%, while concentrations after the perturbation are, in many cases, two orders of magnitude above the original values. Since the latter were based on experimental values, the former may be expected to be out of their physiological range.

The time scale for the lower glycolytic metabolites (BPG to PEP) to achieve their new steady state values is in the minutes range, larger than expected from the perturbation experiment presented in [90][7]. At the same time three of these metabolites exhibit the above mentioned extreme accumulation, suggesting that a smaller change in concentration would be achieved faster.

In contrast to upper glycolysis (G6P to Trio) where level changes are moderate or even negligible (F16P), the lower glycolytic species BPG to PEP increase by some hundred fold. The fact that Trio increases by only 60% while BPG, the next species in the pathway exhibits a roughly 500-fold increase highlights the metabolic upregulation of the reaction *vGAPDH* by its co-metabolites NAD and NADH.

Based on the above, it is suggestive to view upregulation of *vGAPDH* and downregulation of *vPYK* by their respective pairs of cofactors as the main cause for the accumulation of the species between them to levels necessary for these reactions to achieve their new steady state.

[7] We note that, in contrast to our *in silico* experiment, the control parameter, glucose, in this study did not result in a new constant external condition, rather in a high but depleting external glucose level, blurring somewhat the estimation of the time scale of regulation.

Notably, the ATP-dependent reaction $vPGK$ does not seem to exhibit a notable effect on this, reflecting that the concentration change of the adenosine co-metabolites does not alter its potential of ATP generation to a perceivable extent.

Although the model's reaction to oxygen increase should be viewed only as a qualitative approximation to expected behaviour of central metabolism, it seems to be in line with some expectations, enabling us to elucidate on their mechanism. On the other hand, producing qualitatively incorrect behaviour helps to pinpoint additional mechanisms to be represented in future versions of the model.

An example is the examination of the above mentioned non-physiological accumulation of the lower glycolytic species. The following allosteric regulatory links are likely to prevent this in reality, (c.f. Fig. 4.1), since they mediate information between the lower and the upper glycolysis. The lower glycolytic enzyme PYK is activated by F16bP, a metabolite, easily accumulated in simulations due to the irreversibility of PFK, which, in turn is inhibited by PEP, an interaction mediated by F26bP, as discussed in Chapter 3. Including this latter regulatory element may be expected to allow us to tune the steady state levels of the lower glycolytic metabolites by adjusting the corresponding parameter(s), while including PYK inhibition prevents accumulation of F16P in certain scenarios.

The model exhibits a flux redistribution towards TCA-cycle and respiration. At the same time, it exhibits reduction in the production of ethanol and glycerol. Remarkably, the mechanism for the latter is not based on allosteric regulatory elements, but purely on cofactors. The underlying scheme is remarkably simple: upon introduction of a new, highly oxidative agent (oxygen), the reduced form (NADH) of the oxidative power carrier is depleted by, and hence channelled towards the corresponding reaction, consequently, away from competing reactions involving weaker oxidising agents (ethanol and glycerol production) which will therefore experience 'downregulation' due to depletion of their shared substrate. The process is reversible: upon depletion of oxygen, a flux redistribution takes place directing NADH towards other oxidising reactions. The effect is caused by the physiochemical properties of the oxidising agents - lastly, it is simply the higher electronegativity of Oxygen - mediated via a long chain of events - driving the carbon flux towards the respiratory chain. This mechanism has been associated with the Pasteur-effect for a long time [22].

The above argument relies on the assumption that an increased need for other substrates for the involved redox reactions can be satisfied against a chemical potential small enough for the involved oxidising agents to control the flux distribution.

Remarkably, the above mechanism of metabolic regulation is driven by physiochemical properties of the involved reactants and general principles underlying chemical kinetics, rather than regulatory elements, i.e. the latter are not absolutely necessary for it to work. Hence, it is plausible that such regulatory mechanisms for flux redistribution belong to the most ancient layers of the evolution of metabolism. Of course, such processes can further be optimised by regulatory interactions. For example, it may be assumed that (at least for certain parameter sets) the flux redistribution towards the TCA-cycle would take place even without the inhibition of the irreversible reaction $vTCA$ by its

product TCAint. However, the explicit upregulation of this reaction by the depletion of TCAint, caused by demand by respiration, enhances this effect.

Notably, this regulatory element in the model is only an exceedingly simplified representation of the regulation of the pyruvate dehydrogenase complex, c.f. Section 3.3.1. It may be assumed that the complex regulatory interaction between intermediates of the TCA-cycle and the 'gateway reaction' to it is involved in the fine tuning of quick adaptation to various changes involving oxygen supply.

Adenosine co-metabolites feature somewhat less prominently in the regulation of the above changes. This is also due to the choice of external reactions in this work: some reactions involving ATP, most notably glucose intake ($vGLT$), are set to a fixed rate. This prevents the model from reproducing the historically known observation connected to the Pasteur effect: the reduction in glucose intake upon higher oxygen provision. However, we can observe depletion of G6P level, due to constant supply, but metabolic upregulation of HK by higher ATP level. This makes it likely that making glucose concentration an external concentration, rather than having a constant influx, would result in an *increase* of the glycolytic flux. This is in line with the reported behaviour of the Teusink model, according to which higher ATP concentration resulted in an increase of the glycolytic flux [87]. Hence, the model lacks an important regulatory link to reproduce the observed behaviour.

The mechanism currently considered to be responsible for limiting influx into glycolysis in such scenarios is the inhibition of HK by trehalose 6-phosphate [40], [29], [92]. Thus, reproduction of the behaviour of the glycolytic flux during ATP level changes is likely to require the representation of the trehalose biosynthetic pathway.

4.3.3 Discussion

In this chapter we presented a methodology to compare multiple metabolic states of a cell, testing the consistency of a significant amount of experimental information. We applied this approach to compare anaerobic and aerobic metabolic states of *S. cerevisiae* in continuous chemostat conditions.

The kinetic modelling of central carbon metabolism, especially glycolysis, has decades of history by now. While more theoretical studies may be immensely valuable to identify qualitative properties of the system, the biological community - influencing that of systems biology - has been most interested in direct comparison with experimental data. Availability of data has naturally improved during this time, it has become possible to obtain quantitative, rather than qualitative information about various aspects of cellular processes. With the advent of genome scale '-omics' techniques, the amount of information now often surpasses the ability to process it - even if this often involves a tradeoff regarding accuracy.

This has inspired theoretical studies with the aim of more direct comparison to experimental data. Studies using kinetic modelling focusing on experimental data include [67], [68], [41], [71], and [43], to name only a few selected publications studying various aspects of central carbon metabolism. Such kinetic studies focus mostly on regulation within the metabolic level. Perhaps the most impressive example of a kinetic study of metabolic dynamics on the time scale of seconds is found in [71]. Investigating metabolite dynamics on the sub-minute timescale is

connected to the assumption that the direct effect of changes in metabolic levels is the main driving factor on this timescale. Thus, studying flux redistribution connected to interactions within the metabolic reaction network itself - termed 'metabolic regulation' in the present work - is not new.

One the other hand, arguably the most experiment-based kinetic model describing central carbon metabolism was constructed by Teusink *et al.*, [87], who studied a single steady state of this model, testing the consistency of flux, concentration, and various enzyme kinetics data, to a large extent generated within the study.

In a sense, our work may be viewed as a continuation of Teusink's effort: we used a similar set of data and compared two metabolic states - associated with two steady states corresponding to two parameter sets of the model.

We demonstrated that the Teusink model - more specifically the kinetic expressions therein - is capable to satisfactorily reproduce, not only the anaerobic, but also the aerobic data from our datasets. This is notable, since these data were based on a different yeast strain cultured under different (glucose limited) conditions compared to the data used for constructing the model. This consistency may be interpreted as a hint that the involved enzymes are entities with sufficiently stable kinetic properties across the above variations.

The presented methodology is not limited to handling only two states - this only mirrors our focus on elucidating the basic physiological differences between anaerobic and aerobic metabolism, as well as the lack of literature data for microaerobic yeast chemostat cultures.

Comparing multiple metabolic states of the cell corresponding to different conditions allowed us to learn about the effects of non-metabolic regulatory levels on enzyme capacities, lumped together under the term 'hierarchic regulation' (after [73]). We are not aware of a published kinetic modelling study of central carbon metabolism using this approach in order to derive statements about hierarchic regulation, such as the prediction of enzyme capacity fold changes between metabolic states. Importantly, these quantities are rather accessible to experimental verification, even though the extent to which *in vitro* enzyme capacities can be measured remains to be elucidated. Comparison with enzyme level datasets will certainly elucidate further aspects of the regulation of central carbon metabolism. Including data on enzyme activities or enzyme quantities in the presented analysis may be expected to lead to a refinement of the model and information about further regulatory levels.

The amount of experiment based information contained in the present work - culminating in Figure 4.8 - is high compared to similar studies. It contains a detailed account of flux distribution, a sufficiently complete set of concentration data, and a significant amount of information encoded in the kinetic expressions of the Teusink model.

It is a notable result that, not only there exists a set of enzyme capacities consistent with this amount of information, but it is sufficiently well defined to be usefully interpreted. This is not changed by the fact that the model is underdetermined, meaning that the values for some parameters are not well defined by the model's assumptions (c.f. Section 3.4). We restrained from setting these parameter valuess 'by hand', not only since this was not necessary for the presented analysis, but also because their ambiguity within the search range

represents potentially valuable information about the model.

Consistency with transcript data could not be achieved. We interpret this as another hint that level change of a transcript alone does not suffice to predict quantitative changes in the corresponding enzyme's capacities, as reported before [5].

We were sceptical regarding experimental values on cofactor concentrations. While such data are available on ATP-ADP and NADH-NAD concentrations, we felt that the absolute values are less reliable than preferred for quantities that central in the model. The difficulties arising in experiments aiming to determine their absolute concentration range from quick temporal change (c.f. Fig. 9 in [90]), probable large differences between compartments, and distinguishing between free and enzmye-bound forms ([10]). Especially NAD-NADH concentrations are far from being even approximately mapped.

This motivated us to focus on the ratio of cofactor-pair concentrations which we regarded as a more substantial property of the model. However, since both cofactor pairs are determined by a single respective dynamic variable in the model, defining a nominal range for their respective ratios was equivalent to restrict their concentrations in certain bounds.

With values near 1.5, ATP/ADP ratio is somewhat lower than current estimates ([87], [90]) in the presented simulated steady states. Adjusting nominal range during parameter estimation helped to achieve this value, since previous fits had even lower ATP/ADP ratio.

No explicit compartmentalisation is introduced in the model. Introducing different species for cytosol and the mitochondria would inevitably add to the complexity of the model - which translates into an even higher number of parameters - even if transport processes and the various shuttling mechanisms for the nicotinamide species are greatly simplified. At the same time, concentration values in cytosol and mitochondria are not measured in most experimental studies yet.

However, to a certain extent, concentration differences between cytosol and mitochondria can be regarded as encoded in the values of parameters like K_m and K_i of the relevant reactions, since these define effective concentrations. In this model, this is complemented by the fact that all equations used to describe reactions in the mitochondria are highly simplified phenomenological descriptions, such that their parameters are not directly comparable to measured quantities.

Using different K_m values for a metabolite appearing in different compartments amounts to defining a fixed factor for its concentration difference between these compartments. Wile this might be satisfactory for modelling steady states, if each parameter might be fitted separately for each state, it is expected to become a crude approximation when considering time courses.

No new regulatory elements were introduced in the model, apart from the inhibition of $vTCA$ by its product TCAint, in order to avoid product accumulation. The main reason for this was the fact that ambiguity in the corresponding allosteric parameters would, within the presented steady state approach, cause ambiguity in the $V_{\max}$ values, inevitably complicating the analysis. While literature data on some allosteric parameters exist, their uncertainty is rather high.

Whether the failure of the model to reproduce the concentration of the species Triose in the anaerobic experiment is due to a missing regulatory link (e.g. the effect of ATP on G3PDH) remains to be explored.

On the other hand, the presented *in silico* perturbation experiment suggests that a high amount of information about allosteric (or other) parameters may be derived from the requirement that similar *in silico* experiments produce realistic results.

In the case of yeast central carbon metabolism, it may be natural to require that a kinetic model should react in a (qualitatively) realistic way to perturbation experiments involving a glucose pulse, oxygen level change, and aerobic feeding on ethanol. This approach assumes that the most relevant processes influencing these phenomena are contained in the model or can be added relatively easily.

However, it should be pointed out that, while flux data are rather important to constraint a kinetic model, high quality data resulting in a detailed stoichiometric account, such as presented here (based on the efforts in [47]), is only possible to generate for steady state cultures at this point. Hence, while metabolic data on the sub-minute scale allow a fascinating look into the kinetics of metabolism, generating flux data on a similar timescale is, if at all, only possible at the cost of quality, at present.

Drop in ethanol production upon oxygen level increase could be qualitatively reproduced by the model, and further adjusting of parameters may enhance this effect to produce a more substantial change. However, the reported zero flux through ADH during aerobic metabolism, together with non-zero level of its substrate, acetaldehyde (AcA, c.f. flux through *v24* and *v17* in Fig. 4.5) may still be surprising for the following reason. Being presented with non-zero concentrations of its substrates and with zero concentration of one of its products, the reversible model reaction vADH would inevitably exhibit a non-zero rate. Representing various kinetic properties of various members of this isoenzyme family (ADH1 - 7) would not change this fact. There is no known allosteric regulatory link to explain the apparent inactivity during aerobiosis, since it seems unlikely that the enzyme ADH would be not present during aerobic conditions. Not only is there a clear transcriptional activity of ADH isoenzymes in aerobic cultures [69], but lack of this enzyme would render the cell probably unable to cope with sudden oxygen cutoff.

However, the reaction is also regulated by phosphorylation events which, in turn, are influenced by cAMP level [64]. This example highlights the potential importance of a regulatory layer, not included in this analysis - nor in other enzyme kinetic studies of central carbon metabolism of which we are aware.

Among the methods often applied to kinetic models is metabolic control analysis (MCA), a perturbative approach to systems of the form of Eq. 4.1. Among the reasons why this approach was not included in the present analysis, most important is the fact that the author is only aware of implementations focusing on so-called local properties of a single steady state (c.f. [36] [24], [26], [50], [85], [59], and [13]). While control coefficients provide valuable information about steady states of kinetic models (among others), their standard interpretation does not match the main focus of this work: analysing changes *between* different steady states (corresponding to different parameter sets).

Hierarchic regulation - two complementary approaches

Our approach to quantify contributions from the hierarchic and the metabolic regulatory levels is related to that applied in [73], making Table 4.5 a simulation-based analogue to Table 2 in that publication. While both approaches choose measured flux through an enzyme as the quantity whose regulation is under consideration, the principal difference is the following: in [73], the calculation is based on measurement of the enzyme capacity change which associated with hierarchic contribution. That fraction of the measured flux change which is not explained by this, is assumed to be contributed from metabolic regulation (formalised in the 'summation theorem' formula in the paper, see also Section 3.4).

The approach presented here is complementary: we used experimental data on fluxes and metabolic concentrations as well as further information contained in the Teusink model. Finding $V_{\max}$ values consistent with this information can be restated as the assumption that the fraction of the flux change which is *not* explained by the data - i.e. by metabolic regulation - can solely be accounted to changes in $V_{\max}$ values - i.e. to hierarchic regulation.

It follows from the above that, under the assumption that the experiment in [73] measures the extent of hierarchic regulation correctly, the approach in the paper *defines* metabolic regulation, as the source of the flux change contribution *not* accounted to $V_{\max}$ changes.

In terms of this definition, it seems likely that the approach presented in this chapter is more likely to overestimate the effect of hierarchic regulation, rather than underestimate it. This statement is implied by the assumption that non-hierarchic regulatory elements - such as allosteric interactions and covalent modifications - tend to exert their regulatory effect towards the same flux redistribution as hierarchic regulation, when comparing the two experimental conditions under consideration.

We consider two special cases which seem in line with the above assumption within our experiment. The inhibition of the enzyme PFK by PEP was not included in the model. Since PEP increases four-fold from the anaerobic to the aerobic condition, including this link would enable the kinetics used for the model reaction *vPFK* to produce the same concentration and flux values as in the present model with less decrease in the parameter `vPFK_Vmax`.

Regarding the reaction associated with ADH, we argued above that it seems unreasonable to assume total lack of the protein to explain zero flux in the aerobic cultures. This argument was based on the general assumption that central carbon metabolism must react to certain environmental changes on a time scale of seconds in order to fulfil certain essential functions, such as keeping a minimal ATP production or prohibiting the harmful accumulation of metabolites.

The above argument offers an interpretation for the case that $V_{\max}$ changes predicted by our approach would prove systematically different from corresponding measured quantities. In certain cases, it may indicate the necessity of new regulatory elements being taken into account in order to explain flux redistribution, while in others, it can potentially indicate unknown regulatory elements.

Chapter 5

Summary, Conclusions, Future Directions

5.1 Summary

In this work, we aimed to elucidate central carbon metabolism focusing on the aspect of regulation, especially by separating two regulatory levels: metabolic regulation, associated with direct interactions of metabolites and enzymes, and hierarchic regulation, associated with enzyme level change via regulation of *de novo* enzyme production. Our investigations were largely based on the analysis of three datasets from glucose limited continuous cultures of *S. cerevisiae* with five different oxygen provision levels ranging from anaerobic to highly aerobic. These datasets contained data on transcript levels, intracellular metabolite levels, and intracellular flux distributions, respectively. The experiments were performed by collaborators in the VTT Research Centre of Finland.

Part of our effort was the construction of a kinetic model describing key parts of yeast central carbon metabolism under the given conditions, and test, to which extent it can reproduce experimental data, obtaining predictions for enzyme capacity fold changes between anaerobic and aerobic steady states. In this, we leaned on an existing kinetic model, based on experimental data, hence this part of our work can be interpreted as the use of and contribution to the accumulation of biochemical knowledge encoded in the original model.

Extracellular conditions on the macroscopic scale were investigated in Chapter 2. This was inspired by the perceived lack of clarity regarding an important aspect: concentration of glucose, the limiting nutrient and main carbon source in these cultures.

The outcome of this theoretical analysis was

- (*i*) Characterisation of the selection pressure in a chemostat culture, as selecting for cells (or metabolic states) which produce the growth rate, defined by the pre-set dilution rate, with lower external concentration of the limiting nutrient.

- (*ii*) Residual concentration of the limiting nutrient in a chemostat culture

is lower in more advantageous conditions or equivalently, an advantageous condition may be defined as one, under which the pre-set growth rate is produced with lower concentration of the limiting nutrient.

The main assumptions behind these results is that cells in such cultures only receive information from their microenvironment, and changes in the latter are fully characterised by concentrations changes. The relationship between the concentration of the limiting nutrient and the cells growth rate is assumed to be monotonous. If the relationship between the limiting nutrient's external concentration and its specific consumption rate at the fixed growth rate is monotonous, 'residual concentration' in the above can be replaced by specific consumption of the limiting nutrient.

As a consequence (*ii*) characterises selection pressure by stating an objective function in chemostats as long as the above assumptions are met. We discussed possible differences in the character of selection pressure in other continuous cultures and the possibility of cultivations with custom-designed objective function.

Flux regulation on the scale of individual enzymes was investigated for selected reactions, and the phosphofructokinase - fructobisphosphatase (PFK - FBP) unit in Chapter 3. This analysis was based on the attempt to reproduce flux changes through these reactions, using enzyme kinetic expressions with inputs from the three aforementioned datasets. The notion of hierarchic and metabolic regulation was based on earlier work of the Westerhoff group.

Instead of using the original mathematical formulation of this notion in terms of the flux ratio between steady states, defining hierarchic and metabolic coefficients (ρ_h, ρ_m) , we reformulated it in terms of the flux difference, enabling us to use quantities linear in the flux change. The resulting alternative coefficients (R_M and R_H) were interpreted as corresponding to normalised contributions to flux change from the metabolic and hierarchic regulatory levels, respectively. The two kinds of coefficients were compared for a number of hypothetical scenarios. To quantify the extent to which flux data was reproduced, we defined the flux change of 'unknown' origin (R_{unk}).

Hierarchic regulation was associated with transcript level changes in this analysis. This required assumptions on the relationship between changes in transcript levels and corresponding enzyme capacities. Instead of assuming fold changes of the two to be equal, we developed a less stringent assumption, based on a presented model of the regulation of *de novo* protein production. In this approach, the assumption of 'concerted regulation' of various regulatory levels in this process replaces the above assumption. Mathematically, it results in a one parameter formula, in which a parameter (`trexp`) is allowed to vary. We deduced that, if the assumption of 'concerted regulation' holds, the parameter is expected to be close to the number of regulatory levels. The aim of this model was to enable us to formulate a 'null-hypothesis' for a simplicistic view of multi-level regulation of enzyme activity.

The subsequent analysis of the enzymatic reactions associated with fumarase (Fum), phosphoglucose isomerase (PGI), pyruvate dehydrogenase complex (PDHc), and the PFK - FBP unit revealed differing levels of involvement of hierarchic and metabolic regulation in changing the flux through the enzyme. The extent to which the flux change between the five oxygenation conditions could be reproduced, varied greatly, showing that transcription level changes

could not be used as an estimator for hierarchic regulation for two of these four reactions. Notably, no relationship seemed to be present between the complexity of regulation of the reaction, and the extent to which flux changes could be reproduced. Flux changes through PGI and PDHc could not reproduced, while flux through Fum and PFK - FBP was found consistent with the assumption of correspondence between transcription data and enzyme activity.

For the PFK - FBP unit, we combined transcription and concentration data with kinetic expressions (for PFK and FBP), and algebraic constraints, the latter stemming from the steady state assumption for the F26bP branch (PFK2, FBP2). This model showed notable versatility when reducing hierarchic regulation: it could well reproduce measured flux change for a range of `trexp` values, including zero. A possible interpretation of this result is that the metabolic level regulatory network of this set of enzymes is capable of achieving the measured flux change both with and without an active hierarchic regulatory level. Activity of the reverse reaction, FBP, facilitated this versatility. This challenges the opinion, according to which this enzyme's only role is during gluconeogenesis, and suggests the possibility that it plays a role in the control of flux, or maintaining its constance against sudden changes in external conditions.

System-level analysis of central carbon metabolism was undertaken in Chapter 3. Using the information on metabolite levels and flux, a kinetic model representing significant parts of central carbon metabolism was constructed. We aimed to construct a model enabling us to study anaerobic and aerobic metabolism. The following sources were used in the process. Flux distribution and metabolite level data was taken or inferred from the datasets based on anaerobic and fully aerobic chemostat experiments. For some intermediate metabolites, literature data were used. For the co-metabolite pairs ATP - ATP and NADH - NADH, we defined nominal ranges for their respective ratios, based on current opinion.

Most kinetic expressions in the model were taken from a pre-existing kinetic model (Teusink model). We chose this model as a basis, since it had been constructed with a focus on the inclusion of experimental data, and hence, contains a considerable amount of biochemical knowledge encoded in the form of the kinetic expressions and the parameter set. Three reactions, representing the respiratory branch of central carbon metabolism, were added to this model.

In order to arrive at feasible flux distributions, constrained metabolic flux balance analysis was performed, using a stoichiometric network, constructed to be consistent with the model's stoichiometry. This was based on experimental data and analysis, published in connection with the flux dataset.

In order to fit the model to the nominal values resulting from the these sources, we varied a number of parameters, notably all V_{max} values. An extensive parameter estimation process resulted in two sets of parameters corresponding to steady states reproducing, the nominal data values of the anaerobic and the fully aerobic conditions. Anaerobic concentration of one species (Trio) was not well reproduced. Notably, V_{max} most glycolytic reactions turned out to be well defined in the sense that they possess clearly separate estimated values for the above two conditions.

The resulting set of enzyme capacities, as well as their fold changes between the two states, are interpreted as predictions of the model with the property

of being *consistent* with all the included information from the above sources, as well as further model assumptions. We suggest to treat the predictions for these enzyme capacity fold changes as upper limits, since it seems that inclusion of further regulatory links in the model would cause some of the predicted fold changes to decrease.

Finally, an *in silico* perturbation experiment, mimicking the sudden introduction of oxygen into the system, was performed. No further modification or fitting of the model was undertaken for this purpose. We found that the model produces some of the expected behaviour, notably a decrease in ethanol and glycerol production. This was identified to originate from the stoichiometric coupling of the relevant reactions to the cofactors NAD-NADH, as it is known to be the case in reality. Hence, the observed flux distribution is based on the basic design of the modelled metabolic network, rather than being an effect of regulatory elements. However, on the quantitative level, the observed flux change was smaller than expected in reality. In addition, a number of metabolites were clearly outside the physiological range after the perturbation. We analysed discrepancies between simulated and expected behaviour, and highlighted necessary modifications for the model.

5.2 Conclusions

Limits of the concept of hierarchic analysis

In this work, we differentiated between two regulatory layers controlling flux through an enzyme. The network of interactions between metabolites and enzymes, termed metabolic regulation, and the change of enzyme activity by the 'rest of the cell', termed hierarchic regulation. Assumed mechanisms for the former involve reactant level changes and allosteric interactions, while the latter is associated with enzyme level changes via various mechanisms regulating *de novo* protein production.

However not all regulatory mechanisms may be classified as belonging to one of the two categories. We discussed the elaborated mechanism resulting in 'product inhibition via phosphorylation events' of the PDH-complex in Chapter 3, and implemented a minimalistic representation of this mechanism in the kinetic model in Chapter 4. We argued in 4.3.3 that flux regulation through ADH is likely to have a significant contribution from phosphorylation events as well.

There are a number of covalent modifications known to influence enzymatic activity, including phosphorylation, nucleotidylation, and ADP-ribosylation. These events are catalysed by 'converter' enzymes, usually each specific for one direction of the 'conversion' process. These mechanisms significantly increase the cell's capabilities to regulate the target enzyme in a versatile way [83]. The system may often be in a quasi-steady-state in which the ratio of active and inactive forms of the target enzyme depends on a number of parameters, such as the respective activities of the converter enzymes for each direction. The potential versatility in regulation is increased, since the converter enzymes themselves may be regulated by a number of mechanisms, including allosteric and hierarchic regulation (c.f. Section 3.3.1).

This, however, implies that the dissection of flux change into contributions from 'metabolic' and 'hierarchic' levels of regulation is not clearly captured by

analysis of the target enzyme. Namely, change in the target enzyme's activity by covalent modification events may be sourced back to either allosteric or hierarchic regulation of the converter enzymes' activity (c.f. Section 3.3.1). From a more general point of view, this is an implication of a principal difference between 'pure metabolic networks' and systems involving cascades of covalent modifications: a strict differentiation between 'reactant' and 'enzyme' is only possible in the former. In addition, phosphorylation events can be induced by signalling cascades, which may, therefore, affect metabolism directly, i.e. without involvement of the (slower) hierarchic level regulation. For example, the important mid-glycolytic enzyme PFK is effected by HOG pathway activity: phosphorylation of PFK2 changes the level of F2,6bP, an effector molecule of PFK, (c.f. Section 3.3.2), [20], [21].

We are not aware of enzyme-level models of yeast central carbon metabolism which include this regulatory level (disregarding the exceedingly simplified representation used in the model presented here). The role of covalent modification in the regulation of yeast central carbon metabolism is presumably important in many scenarios. A useful compilation of known interactions in [64] is shown in Fig. 5.1. We assume that representation of this regulatory level in future models will be necessary for understanding certain classes of scenarios. However, to our knowledge, availability of experimental techniques continues to represent a major limitation here, since common measurement techniques for enzyme activities - such as used in [53], [73] for the calculation of hierarchic coefficients - and enzyme levels are often not able to take phosphorylation levels into account.

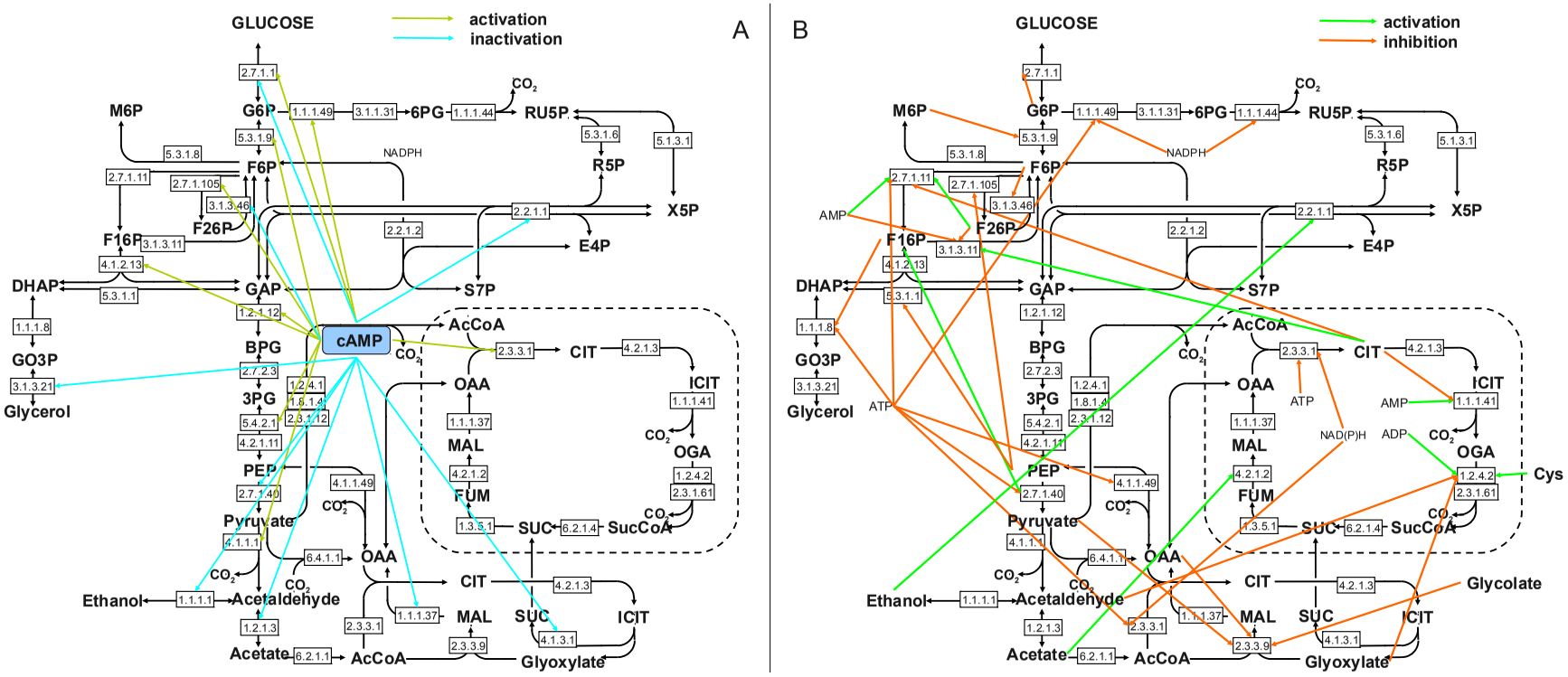

Figure 5.1: Summary of covalent (Panel A) and allosteric (Panel B) modifications of enzyme activities in yeast central metabolism. In most cases, cAMP effects the activity of protein kinases or protein phosphatases, the latter performing covalent modifications of the target enzyme. Based on the databases BRENDA and YPD. (Modified from a Figure in [64])

Timescales of regulatory levels

In understanding a system containing various interacting components, assessing characteristic timescales for the dynamics of these is often a powerful tool. In the context of metabolism, this approach has been used to differentiate between interactions within the metabolic network, as we have discussed in Section 4.1.

Here, we regard the experimental setup - such as a continuous cell culture - as a systems, components of which range over several orders of magnitudes regarding size, volume, and timescale of dynamics. Cellular compartments of a yeast cell are in the femtoliter range, while the working volume of a typical chemostat is in the liter range.

In a chemical reaction network, the characteristic time scale for the dynamics of concentration may be estimated by assessing the possible range of concentration/flux ratios for the relevant species. Given the low concentration and high consumption rate of the limiting nutrient in a chemostat culture, the relevant timescales for the dynamics of its external (residual) concentration are often comparable to that of intracellular processes (for example, when disrupting the glucose influx by removing the sample from the bioreactor).

A time scale assessment for some components of interest, resulted in the following table:

Time scale	process, comment (reference)
sub-second	certain metabolic levels [19]
second	internal metabolites [90] & residual glucose (Chemostat) [74]
minute	phosphorylation [37] & external metabolites (Chemostat) [90]
10s of minutes	enzymatic capacity [9] & phosphorylation [21] & Chemostat [9]
hours	protein level [70] & external metabolites (Chemostat) [97]

For each component, timescales are indicated at which it displays active dynamics after a perturbation, i.e. it has already reacted but not yet attained a new steady state. It is remarkable that components of regulation of metabolism are separated so clearly regarding their characteristic timescales. After a sudden perturbation, internal metabolite levels are believed to attain a new steady state within ca. 100 seconds, well before further changes due to other regulatory events start effecting them on a slower timescale.

We are not aware of an *a priori* reason for the various regulatory levels not to have 'overlapping' time scales. This would result in a lack of clearly perceivable steady states after a perturbation. On the other hand, experiments involving the timescale of seconds ([91], [90], [72], [71]), in which timescale separation was observed, require great effort and have been performed only under certain experimental conditions.

For covalent modification events, concentration of both target and converter enzymes may vary in a rather wide range, hence we assume that the dynamics of phosphorylation events covers a broad range of timescales. While signalling cascades are known to be able to process information within a minute [95], the above cited study [21] suggests a characteristic timescale in the minute to hour range for the dynamic regime.

Regarding the reaction catalysed by ADH in yeast (c.f. Section 4.3.3), it can be argued that its activation by phosphorylation events should be possible within a the time span which the cell is able to spend without regeneration of ATP from ADP - otherwise, a sudden oxygen cutoff would be lethal.

Cooperation of regulatory levels

The main players of central carbon metabolism - the enzymes - are largely identified, to a certain extend characterised, especially for model organisms of interest, such as *S. cerevisiae* or *H. sapiens*. We seem to have gathered enough information to start understanding the dynamics within certain regulatory networks. Kinetic models of metabolism, a major focus of this work, represent our current understanding of the metabolic regulatory layer.

We have not been concerned with the details of pre-translational regulatory mechanisms in this work, apart from a rather simplicistic treatment in Section 3.2.2 with the aim of defining a 'null hypothesis of simple regulation'. Some components of these regulatory layers, such as the regulation of transcriptional networks [35], [31], have been studied for some time, resulting in a certain level of understanding - even though approaches aiming at quantitative understanding are in a less mature stage than for kinetic models of metabolism. Other regulatory mechanisms, such as microRNAs [4], have only been recognised recently.

We implicitly assumed that these regulatory levels are 'slow', i.e. the characteristic timescale of their dynamics after an external perturbation is above that of minutes.

However, understanding the regulation of metabolism cannot be complete without the understanding of how the various regulatory layers work together, and this is a major focus of present day research [60], [7], [31]. Is it possible to identify any general principles, and further, use them to aid model construction? Before attempting to answer this questions, we discuss the two relevant examples we encountered in this work.

The model of the PFK-FBP regulatory system, constructed on the basis of experimental data as discussed in Section 3.3.2, turned out to be surprisingly flexible regarding the extent of hierarchic regulation, when attempting to reproduce measured flux changes. The model was capable of reproducing the measured metabolite concentrations and fluxes with, as well as without amplifying activity fold changes of the involved enzymes. Notably, the PFK-substrate F6P is known to be in quasi-equilibrium with G6P, while the latter is the starting point for major pathways, hence, its concentration directly effects their flux. Consequently, regulating F6P concentration has a potentially large impact on the flux through these pathways, depending on the presence of further regulatory mechanisms.

On the pathway level, we discussed metabolic versus hierarchic regulation with respect to necessary flux changes upon sudden changes in the availability of nutrition or oxygen (c.f. Section 4.3). The underlying assumption was that certain regulatory events must take place fast (on the timescale of seconds) for basic metabolic functions to be provided. Scenarios requiring fast regulation include sudden changes in nutrient or in oxygen provision, as well as various other stress scenarios. This excludes the possibility of hierarchic regulation to be alone in charge of counterbalancing such events. The metabolic model presented

in Chapter 4 was shown to produce flux redistribution regulated by the metabolic regulatory level upon sudden introduction of oxygen in the anaerobic state.

Above, we have discussed the apparent separation of timescales of the characteristic dynamics of the regulatory levels. A possible explanation for this separation may be based on an evolutionary argument, at least applicable to single-cell organisms.

It seems natural to divide metabolic regulatory levels according to their characteristic timescale. Fast regulation may be expected to have evolved to be able to produce an appropriate response to the *widest* possible range of sudden perturbations of external conditions. On the other hand, 'fine tuning' towards optimal resource usage may be expected to have lower priority on this timescale.

In contrast, regulatory levels acting on slower timescales may be expected to have evolved to provide a metabolic state characterised by optimum principles, such as optimal usage of resources. Hence, selection pressure associated with steady states, such as the one discussed for chemostat cultivations, may be expected to determine the metabolic state with regard to this longer timescale.

Importantly, the above notion does not characterise the involved regulatory components individually, it should rather be understood as a characterisation of certain requirements regarding their *cooperation*.

The above view offers the possibility to interpret the timescale separation observed in the regulation of metabolism, as an answer to a basic question in system design: optimisation versus specialisation. Interestingly, the above considerations offer a systemic definition of 'stress' as a metabolic state in which the fast regulatory level is outside of its 'operational conditions', i.e. not capable to upkeep the normally required conditions for the rest of the cell, resulting in the necessity of the slow regulatory level to react fast, if harm to the cell is to be avoided.

Within the framework of computational modelling, the above considerations may be formulated as optimisation principles, and potentially be implemented in the form of simulations of 'competing metabolism models' in a changing environment.

Recently, the integration of kinetic modelling into the experimental community's view seems to have lead to a drop in the perceived role of the slower, hierarchic regulatory levels in flux redistribution. However, for example in the case of the impressive double-perturbation experiment reported in [9], the conclusion about the minor role of this regulatory level seems to have been made without the above considerations. The time period of the experiment after the perturbation was two hours, and it was terminated before all enzyme activity levels (which we assume to be regulated by hierarchic regulation) reached a new steady state (c.f. Enolase, Fig. 5 in [9]).

A suggested experiment

Quantitative modelling is, in principle capable of predictions, however this is offset in praxis by the amount of information needed to construct the model. Data on internal metabolites and fluxes represent key inputs for model identification, i.e. they help in deciding between a class of *a priory* equally good models. However, they are expensive in terms of equipment and technology.

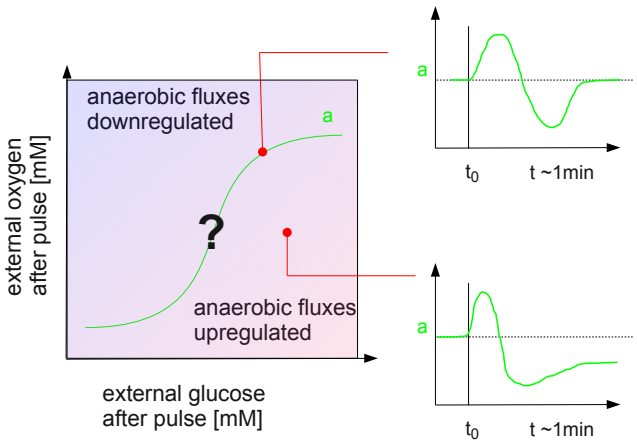

Figure 5.2: A proposed series of double-perturbation experiments to find balance-curves for antagonistically regulated quantities. a denotes a quantity - such as glucose intake or ethanol production - which is antagonistically effected by external oxygen and glucose increase when introduced in a glucose limited continuous yeast culture under microaerobic conditions.
The axes indicate the level at which glucose, and oxygen are kept constant after the perturbation. Hence, each point within the axes defines a double-perturbation experiment with a given level change of these two external species.
It is to be tested whether points exist at which the post-perturbation steady state of a is the same as prior to the perturbation, as indicated in the inlay in the upper-right corner. If - as indicated in the figure - these points give a functional relationship between the extent of glucose pulse and oxygen pulse, they can serve as useful characteristic of the regulatory system. The indicated timescale (1 min) selects metabolic regulation to be assessed.

In the following, we suggest a class of experiments which may result in valuable information about the regulation of central carbon metabolism, even if performed with regard to external fluxes, such as specific consumption or production of a metabolite by the cell. Change in these fluxes is detectable online within a minute [90]. Our suggested approach is to map perturbations which cause a certain variable to exhibit *no* apparent change, when comparing its value prior to the perturbation to that after the new steady state has been attained. This involves combining the following two perturbation experiments, often performed in *S. cerevisiae* cultivations.

Oxygen level increase is known to divert carbon flux away from the branches producing glycerol and ethanol, and towards the respiratory chain. Part of this effect is due to metabolic regulation by the redox cofactor pair NAD-NADH, likely enhanced by other mechanism, such as the above-mentioned regulation of the pyruvate dehydrogenase complex by its dedicated kinase-phosphatase-pair. In addition, oxygen increase causes glucose intake to decline due to the sudden increase of ATP production and regulatory events, believed to effect PFK and HK (via trehalose6-phosphate), among others.

On the other hand, increase in external glucose level results in higher influx of glucose which is, even under aerobic conditions, mostly metabolised anaerobically to produce ethanol (overflow metabolism, [90]).

Hence, for a given cultivation, and for each of the external metabolites, there is a characteristic flux change, corresponding to a given level change in glucose

or oxygen.

Since the above mentioned two perturbations act antagonistically on glucose intake and ethanol production (and possibly on other variables, depending on strain and condition), we may pose the question if is it possible to 'cancel out' a given flux change by applying a specific double perturbation. In other words: can a certain glucose level increase be matched with an oxygen level increase such that the variable under consideration returns to its former value after a short transient? If the answer is positive, a systematic evaluation of the results may result in a graph indicated in Fig. 5.2 for each targeted variable.

As discussed above, for a short period after the perturbation, these changes are predominantly determined by metabolite level interactions (presumably including covalent modifications induced by metabolite level changes, such as cAMP, c.f. Fig. 5.1). Hence, the value of the selected variable approximately 100-200 seconds after the double-perturbation is likely to be the new steady state value determined by the regulatory levels associated with this timescale (c.f. [90]).

Should the experiment reveal no steady state in this time window, a possible interpretation would be that time scale separation of the involved regulatory levels is not as clear as generally assumed, based on experiments like [90].

5.3 Speculations

On consistency with physics

Compared to physics, generally assumed to be the most formalised of the exact sciences, biology lacks general quantitative framework theories which are assumed to be valid over a large range of phenomena and, notably, severely restrict the form of eligible models for a certain system. Theories, such as electrodynamics or mechanics possess their rather specific sets of equations, which play an important role in restricting system models to such, which are compatible with them.

Due to the high level of formalisation, this process often does not even need to be stated explicitly - for simpler systems the appropriate equations can be constructed directly, from the general forms, sometimes even in a nearly unique way: the equations describing a pendulum, may be written down unambiguously, once the framework theory within which to operate - classical or quantum mechanics, electrodynamics - is decided.

In contrast, few framework theories exists, which would play a similarly important role for biology. (The obvious extension of physical theories to object which happen to be 'alive', as in biomechanics, seems to play a lesser role in understanding biological problems). As a result, models describing biological systems are much less *a priory* restricted, causing model identification to be a major issue.

Arguably, at least two theories may aspire to a status within biology, similar to the above mentioned framework theories: evolution and thermodynamics. They are assumed to be of general validity, and are, in principle, capable of restricting models by the requirement of consistency. For example, a model of a metabolic pathway which violates thermodynamics has - in principle - to be regarded as wrong. Similarly, a scenario which cannot be fit into the framework of evolution, will be rejected, for example the existence of a hypothetical epidemic

with 100% lethality.

YET, ADHERENCE to these two theories is presently not aimed for in typical kinetic models of metabolism, including the one presented in this work. It is - in principle - possible to reject parameter sets for a kinetic model, rendering it inconsistent with thermodynamics [8, Borger], [10, Canelas *et al.*]. In praxis, however, this is seldom taken into account in the construction of biochemical pathway models. It seems, among the major reasons is, once again, the lack of experimental methods to determine thermodynamic properties of the relevant processes *in vivo*.

On consistency with evolution

Based on the above considerations, we may formulate certain desired properties for an ideal kinetic model of metabolism as follows:

- (i) Consistency with thermodynamics. This requires improvement of our knowledge about free energy changes associated with various metabolic processes under *in vivo* conditions.
- (ii) Existence of fast regulation which should ensure that known major obligations of metabolism - such as keeping cofactor and reactant levels within feasible ranges - are fulfilled at all times, with regard to a wide range of stress perturbations of external conditions representable in the model. For *S. cerevisiae*, relevant scenarios include sudden shifts from aerobic to anaerobic condition, from glucose feeding to gluconeogenesis, and switching between major carbon sources. Non-metabolic perturbations include oxidative or osmotic stress - modelling these naturally requires representation of relevant cellular components in the model.
- (iii) Long term adaptation. The model should be able to reproduce optimisation of relevant quantities - such as given growth rate with minimal glucose flux in case of a chemostat cultivation - to the extent, known from experiments. Again, formulation of many optimisation principles requires the representation of certain events outside of central carbon metabolism, for example in order to assign a growth rate to a given metabolic state, when this is required.

A possible approach to model construction with the above requirements involves modelling of evolution under selection pressure. Realistic modelling of bioevolution is a nearly hopeless task, given the complexity of most real existing ecosystems. However, focusing on the comparatively well-understood aspects - such as the metabolism of single-cell organisms - and using pre-existing knowledge - such as that about the ecological niche it occupies - may lead to a better understanding of metabolism, and its regulation in the light of evolution.

Examples for aspects which may be further elucidated with regard to their relationship to selection pressure include the so-called 'turbo design' of glycolysis, [88], [89], or its oscillatory behaviour. This latter phenomenon, previously regarded as an odd 'side product' of certain laboratory conditions [17] may turn out to be the rule, rather than the exception [80].

In an ideal case, the above outlined approach might reproduce aspects which are regarded today as mere facts, thus providing an 'explanation' in the light

of evolution. The ultimate goal of deduction - as formulated for theoretical physics - to explain the principle building blocks for everything with nearly no experimental input, will presumably not be within reach for biological systems for a long time.

Nevertheless, once a considerable amount of quantitative characteristics of a biological field, such as metabolism, can be deduced from the requirement of consistency with physical theories and with evolution, and 'frozen accidents' in the latter can be separated from features arising from natural selection, then theoretical biology will have achieved maturity.

Appendix A

Appendix to Chapter 2

A selection from the transcription (affymetrix) datased, as relevant for this thesis (published in [69])

gene	0% oxygen	0.5% oxygen	1% oxygen	2.8% oxygen	20.9% oxygen
HXK1:	1.3505e+04	1.1542e+04	1.1602e+04	1.0628e+04	7.4638e+03
HXK2:	3.6073e+03	2.0410e+03	2.5013e+03	3.2610e+03	1.5345e+03
GLK1:	6.4137e+03	6.4204e+03	6.2486e+03	4.9109e+03	7.0634e+03
PGI1:	6.6746e+03	5.8718e+03	5.8489e+03	5.7088e+03	6.3717e+03
PFK1:	3.9895e+03	3.3412e+03	3.3127e+03	3.4737e+03	2.4602e+03
PFK2:	4.5542e+03	3.6923e+03	3.7093e+03	3.8204e+03	3.1678e+03
PFK26:	780.1702	899.6081	809.2803	746.3032	1.5404e+03
PFK27:	318.4862	189.4616	203.5313	172.8145	67.4453
FBP1:	63.4244	156.5181	185.4137	291.5665	731.7097
FBP26:	243.9565	326.1265	303.6309	306.7182	439.8271
FBA1:	1.4234e+04	1.4719e+04	1.4537e+04	1.4750e+04	1.5379e+04
GPD1:	1.7589e+03	2.5891e+03	2.4380e+03	1.6699e+03	3.6247e+03
GPD2:	389.5913	201.9018	182.3522	224.7588	317.1473
RHR2:	4.7039e+03	2.6257e+03	2.2949e+03	2.6163e+03	2.7713e+03
HOR2:	1.1992e+03	1.0252e+03	927.9391	558.0686	956.2443
GUT1:	743.9821	3.8097e+03	3.9384e+03	3.4494e+03	3.8971e+03
TDH1:	4.1433e+03	6.2768e+03	6.8143e+03	4.9592e+03	7.1373e+03
TDH3:	2.3049e+04	2.3370e+04	2.3343e+04	2.2918e+04	2.3715e+04
PGK1:	1.2325e+04	1.1797e+04	1.1889e+04	1.0525e+04	1.2251e+04
GPM1:	1.6369e+04	1.4571e+04	1.4768e+04	1.4345e+04	1.5899e+04
ENO1:	1.0290e+04	1.1431e+04	1.1024e+04	7.9131e+03	1.3519e+04
ENO2:	1.4022e+04	1.1166e+04	1.1190e+04	1.0872e+04	1.3474e+04
CDC19:	1.4236e+04	1.2320e+04	1.2055e+04	1.2245e+04	1.3618e+04
PYK2:	549.7716	389.6056	358.1519	336.1659	403.3803
PDC1:	1.2055e+04	1.2391e+04	1.1290e+04	1.0811e+04	1.1882e+04
PDC5:	334.6819	412.7651	388.0064	359.3772	271.3712
PDC6:	204.4989	491.8631	460.5652	274.8636	946.9587
ARO10:	22.9886	73.1132	103.0349	151.0100	239.5289
THI3:	273.5343	397.5374	391.2891	363.0549	388.3693
PYC1:	1.0441e+04	7.4594e+03	7.2248e+03	6.5713e+03	4.0461e+03
PYC2:	1.4792e+03	2.2965e+03	2.1956e+03	2.0970e+03	1.9348e+03
ADH1:	1.7196e+04	1.4761e+04	1.4697e+04	1.4870e+04	8.4274e+03
ADH2:	242.2227	354.8648	443.0525	5.4609e+03	1.5789e+04
ADH3:	7.2239e+03	5.4665e+03	4.9217e+03	4.5971e+03	3.0342e+03
ADH4:	389.5626	389.6140	409.9285	389.7842	307.7876
ADH5:	1.7724e+03	735.1538	654.6271	506.1380	425.8074
ADH6:	2.2932e+03	968.2210	1.1491e+03	1.6777e+03	994.6064
ADH7:	16.3087	17.4191	16.9218	16.1053	16.0399

(turn page)

(continued from last page)

gene name	0% oxygen	0.5% oxygen	1% oxygen	2.8% oxygen	20.9% oxygen
YIA6:	158.7082	223.2772	224.2737	218.3112	220.0259
PDA1:	6.5260e+03	6.2249e+03	6.5801e+03	6.1067e+03	6.2917e+03
PDB1:	2.9374e+03	3.7291e+03	3.9490e+03	3.7272e+03	3.2635e+03
LPD1:	4.7532e+03	7.0674e+03	6.7650e+03	6.2055e+03	5.7597e+03
LAT1:	4.6415e+03	4.5837e+03	4.6092e+03	4.1492e+03	5.4352e+03
PDX1:	676.3640	605.4573	602.4863	614.6312	739.8591
CIT1:	7.1240e+03	9.2908e+03	9.9117e+03	1.0015e+04	7.0200e+03
CIT3:	94.6529	126.3489	139.7151	538.1178	1.9384e+03
ACO1:	5.3042e+03	7.1496e+03	7.5714e+03	8.1567e+03	6.4737e+03
ACO2:	1.7490e+03	1.8342e+03	1.9679e+03	2.0494e+03	1.3937e+03
IDH1:	1.6854e+03	1.9389e+03	2.5293e+03	3.0294e+03	1.8240e+03
IDH2:	2.1865e+03	2.7667e+03	3.3354e+03	3.6882e+03	1.9316e+03
IDP1:	1.4919e+03	1.5496e+03	1.6350e+03	1.4241e+03	1.0845e+03
IDP2:	86.2705	208.4386	306.8240	1.1635e+03	2.8253e+03
IDP3:	114.2046	132.4659	139.9509	226.7480	1.1870e+03
KGD1:	1.6000e+03	3.2014e+03	3.0276e+03	2.6988e+03	2.0754e+03
KGD2:	1.9535e+03	3.9566e+03	4.0979e+03	4.0178e+03	2.7942e+03
LSC1:	2.3259e+03	3.2547e+03	3.4500e+03	4.0024e+03	4.1438e+03
LSC2:	1.7529e+03	2.8366e+03	2.9865e+03	3.2038e+03	3.1211e+03
SDH1:	3.7572e+03	7.6167e+03	7.7838e+03	7.3785e+03	5.0227e+03
SDH2:	4.3786e+03	8.1435e+03	8.5452e+03	7.9696e+03	6.1718e+03
SDH3:	3.8741e+03	6.0066e+03	6.2861e+03	6.4908e+03	4.9670e+03
SDH4:	3.0693e+03	5.5470e+03	6.0546e+03	5.8748e+03	3.3507e+03
FUM1:	3.4403e+03	4.8873e+03	5.2410e+03	5.8009e+03	5.9319e+03
MDH1:	5.7893e+03	8.3640e+03	8.4206e+03	7.4454e+03	6.3575e+03
MDH2:	625.0861	462.3249	447.7491	1.0473e+03	4.0413e+03

Steady state metabolite data reported in the publication [97], received from Dr. Marilyn Wiebe

	G6P					**F6P**				
cultivation	aerobic	2.8% O2	1% O2	0.5% O2	0% O2	aerobic	2.8% O2	1% O2	0.5% O2	0% O2
IDs	(19,20,32,33)	(14)	(23,24,30,31)	(21, 22)	(15,16,28,29)	(19,20,32,33)	(14)	(23,24,30,31)	(21, 22)	(15,16,28,29)
O2 (%)	20,9	2,8	1	0,5	0% O2 (15,16,28,29)	20,9	2,8	1	0,5	0
median	3,024	0,901	2,852	1,496	4,846	0,608	0,155	0,568	0,369	0,869
stdev	0,536	0,618	1,984	0,669	0,988	0,135	0,149	0,338	0,166	0,166
stdev%	18	69	70	45	20	22	96	59	45	19
	2,830087	1,43963675	0,2063766	1,37449209	3,64572047670639	0,584006	0,386485	0,037981	0,284167	0,702194
	2,893022	0,95987346	0,3594784	1,623392943	3,13902243589744	0,818959	0,170455	0,059934	0,368769	0,535969
	2,486795	0,74402916	0,2238694	1,537552333	4,92199764521193	0,660816	0,155038	0,032883	0,372765	0,881083
	2,243961	0,61452732	1,6763124	3,347805495	2,44894924662966	0,599749	0,128497	0,339691	0,777929	0,542724
	2,693638	0,90130893	1,6763124	1,495762393	4,05803925455987	0,653261	0,131367	0,371905	0,392515	0,852994
	2,439179	0,38494532	0,9573557	1,417444178	5,14585384222612	0,451143	0,089932	0,234798	0,267021	1,029545
	3,321	2,22743328	0,2575365	0,97364096	5,35878399829167	0,317485	0,476648	0,052158	0,200716	0,935198
	3,309228		0,1747462	1,41176091	5,15896123016676	0,63678		0,030871	0,295853	0,912699
	3,769608		0,2682846	1,795866704	4,83182946751309	0,431764		0,036275	0,385954	0,856525
	3,757655		1,4045083		4,86075450534168	0,684774		0,298379		0,886728
	3,475305		1,36115758			0,588243		0,288079		
	2,836533		2,4949331			0,299011		0,563887		
	3,023619		5,0476699			0,6596		0,701856		
	4,10931		6,2249471			0,708588		0,763632		
	3,48864		4,3275581			0,653787		0,607887		
	3,404008		4,6517687			0,61571		0,83844		
	2,320128		3,9438269			0,494736		0,859318		
	3,025323		4,2891690			0,527461		0,842906		
			3,9301414					0,779949		
			4,1283975					0,788021		
			3,2083428					0,572591		
			5,2715818					0,785934		
			5,2776539					1,110257		
			3,2160432					0,736534		

	M6P					**T6P**				
cultivation	aerobic	2.8% O2	1% O2	0.5% O2	0% O2	aerobic	2.8% O2	1% O2	0.5% O2	0% O2
IDs	(19,20,32,33)	(14)	(23,24,30,31)	(21, 22)	(15,16,28,29)	(19,20,32,33)	(14)	(23,24,30,31)	(21, 22)	(15,16,28,29)
O2 (%)	20,9	2,8	1	0,5	0,1	20,9	2,8	1	0,5	0,1
median	0,591	0,272	0,791	0,410	1,157	0,131	0,038	0,025	0,041	0,027
stdev	0,115	0,185	0,500	0,203	0,325	0,038	0,009	0,024	0,016	0,018
stdev%	20	68	63	49	28	29	23	97	39	66
	0,63641	0,40651709	0,0452773	0,305056637	0,83085048754063	0,077455	0,04113	0,015387	0,009764	0,033368
	0,818773	0,27784715	0,0652236	0,461363419	0,81152065527066	0,105194	0,031793	0,021628	0,04192	0,03359
	0,695567	0,26122441	0,0384905	0,392525895	1,25343406593407	0,08997	0,042152	0,010461	0,019783	0,071919
	0,647259	0,187047	0,4129763	0,947395427	0,57940126883426	0,071444	0,027544	0,082569	0,041225	0,014044
	0,572486	0,27245566	0,4017605	0,437149636	1,17838025376685	0,126841	0,037538	0,080704	0,059654	0,041978
	0,459474	0,12608409	0,2648452	0,410211866	1,47394430813475	0,10531	0,020033	0,028046	0,057861	0,020248
	0,435024	0,68897174	0,0603088	0,25080311	1,51982689484791	0,111111	0,043393	0,018574	0,036402	0,014175
	0,609614		0,0403950	0,368072508	1,54336030164019	0,137662		0,015319	0,040329	0,016699
	0,483562		0,0503668	0,547626835	1,1351132372665	0,125056		0,011843	0,045082	0,033503
	0,578101		0,4120703		1,10449958147222	0,134538		0,019347		0,01887
	0,502035		0,3941269			0,168786		0,021333		
	0,340304		0,8018383			0,152722		0,030178		
	0,603818		1,2034676			0,161647		0,03438		
	0,72436		1,6123815			0,196274		0,048114		
	0,652899		1,1587401			0,197823		0,031207		
	0,644872		1,2146642			0,165706		0,032079		
	0,525172		0,9542759			0,125078		0,060797		
	0,510046		1,0478218			0,179921		0,057196		
			0,8886104					0,06246		
			1,0041519					0,055041		
			0,8747076					0		
			1,2896975					0,002158		
			1,3358150					0,011658		
			0,7804422					0,002665		

	FBP					**3PG**				
cultivation	aerobic	2.8% O2	1% O2	0.5% O2	0% O2	aerobic	2.8% O2	1% O2	0.5% O2	0% O2
IDs	(19,20,32,33)	(14)	(23,24,30,31)	(21, 22)	(15,16,28,29)	(19,20,32,33)	(14)	(23,24,30,31)	(21, 22)	(15,16,28,29)
O2 (%)	20,9	2,8	1	0,5	0,1	20,9	2,8	1	0,5	0,1
median	0,752	1,148	1,092	1,008	17,119	7,248	3,665	5,457	5,363	3,267
stdev	0,780	0,431	0,600	0,344	3,255	2,171	1,390	2,050	1,917	0,746
stdev%	104	38	55	34	19	30	38	38	36	23
	1,328303	0,75610931	0,2128947	1,109075639	15,2621846929859	8,695348	4,607228	1,058396	3,407683	2,451537
	0,738736	0,72314484	0,2871642	1,350685729	14,3954460561708	10,21185	3,664816	1,995021	9,001331	3,022016
	0,765728	1,60683439	0,2145998	0,575604413	16,4803719194235	10,86542	3,318822	0,605606	3,851864	3,412333
	1,36505	0,80195735	2,3865639	1,66328144	11,4490280299609	12,12582	2,72031	3,363146	7,52427	1,724449
	0,692273	1,14802675	2,2564643	0,941857271	21,9716305392289	7,316171	4,167164	3,245651	5,957491	4,077985
	0,442778	1,16250411	0,3759272	1,008052911	20,121404517087	6,510352	1,465449	5,968628	5,554331	4,258992
	1,977962	1,81999292	0,31116263	0,660896698	16,3912718700675	3,978464	5,80892	1,394207	3,747862	3,393213
	0,634399		0,3816895	0,761567761	20,7341462309001	6,752629		1,630296	3,821622	3,550478
	1,8366		0,38111888	1,148851106	17,7585284491664	3,977198		1,447209	5,363343	2,816343
	1,142943		1,6440973		19,6239579767882	6,48706		5,438312		3,141661
	2,160999		1,05911242			7,220475		5,47629		
	3,318341		1,4617089			5,62276		6,807019		
	0,535127		1,2442391			9,045024		6,387515		
	0,783975		1,48911311			9,145455		6,69272		
	0,457474		1,2981097			8,386743		5,894677		
	0,491664		1,1740317			7,274533		5,970545		
	0,60213		1,0417910			7,10279		5,271985		
	0,508228		1,1282968			6,198123		6,112821		
			0,9191972					5,928264		
			1,11224795					5,985115		
			1,0722904					5,085901		
			1,4766562					5,594543		
			1,3759609					6,034117		
			0,9835700					4,097448		

	Pyruvate					**Fumarate**				
cultivation	aerobic	2.8% O2	1% O2	0.5% O2	0% O2	aerobic	2.8% O2	1% O2	0.5% O2	0% O2
IDs	(19,20,32,33)	(14)	(23,24,30,31)	(21, 22)	(15,16,28,29)	(19,20,32,33)	(14)	(23,24,30,31)	(21, 22)	(15,16,28,29)
O2 (%)	20,9	2,8	1	0,5	0,1	20,9	2,8	1	0,5	0,1
median	0,240	0,701	0,740	0,549	1,388	0,167	0,348	0,311	0,213	0,685
stdev	0,075	0,398	0,675	0,303	0,363	0,050	0,098	0,878	0,069	0,730
stdev%	31	57	91	55	26	30	28	283	32	107
	0,34112	1,30060537	0,1576007	0,53127325	1,61867915780723	0,166885	0,377155	0	0,098805	1,036608
	0,262695	1,1252463	0,1323996	1,237757727	1,8812796905873	0,095799	0,314413	0,058348	0,237813	0,367178
	0,286277	1,0642747	0,0698671	0,535054326	1,20891148740995	0,170356	0,373353	0	0,167043	2,83801
	0,423477	0,38727061	0,2366831	0,871954625	0,97124869817569	0,256091	0,292815	3,495794	0,212308	0,298836
	0,187605	0,66270954	0,4336840	0,680169201	1,96151280761085	0,179017	0,348123	3,129139	0,31297	1,230892
	0,25014	0,22590194	0,2649873	0,549262538	1,58565842522044	0,157097	0,297393	0,312378	0,277241	0,709744
	0,266175	0,70073432	0,0800923	0,218177771	1,56780562986751	0,19566	0,577058	0,056278	0,213135	0,660531
	0,249071		0,0675806	0,317448657	1,12009009017385	0,172116		0,006156	0,205781	0,656672
	0,286486		0,0918991	0,734850005	1,14110793383666	0,204591		0,049653	0,309967	1,058101
	0,257623		0,3481238		0,99954704930194	0,166839		0,304077		0,576938
	0,230899		0,4508605			0,2326		0,315051		
	0,169967		0,6999483			0,229353		0,450025		
	0,140036		2,5673072			0,141265		0,495271		
	0,203702		1,6650122			0,12482		0,56722		
	0,198004		1,6773946			0,097853		0,439645		
	0,178807		1,3030925			0,116244		0,468087		
	0,176387		1,4996518			0,134006		0,373462		
	0,103062		1,01331136			0,078855		0,317006		
			1,1815778					0,308921		
			0,83416118					0,269579		
			1,7019126					0,284977		
			1,0026032					0,33936		
			0,9871064					0,257473		
			0,7797821					0,149143		

	pentoseP					**PEP**				
cultivation	aerobic	2.8% O2	1% O2	0.5% O2	0% O2	aerobic	2.8% O2	1% O2	0.5% O2	0% O2
IDs	(19,20,32,33)	(14)	(23,24,30,31)	(21, 22)	(15,16,28,29)	(19,20,32,33)	(14)	(23,24,30,31)	(21, 22)	(15,16,28,29)
O2 (%)	20,9	2,8	1	0,5	0,1	20,9	2,8	1	0,5	0,1
median	0,162	0,284	0,150	0,124	0,465	3,122	1,112	1,776	1,791	0,857
stdev	0,113	0,066	0,080	0,089	0,368	1,242	0,501	0,741	0,797	0,399
stdev%	70	23	53	72	79	40	45	42	45	47
	0,322948	0,27916606	0,0326911	0,253369789	1,16911814214273	3,356055	1,338046	0,328071	0,858727	0,214831
	0,348985	0,28408503	0,05440112	0,273686483	0,77839546412993	4,229101	1,112206	0,534342	3,365273	0,26145
	0,367463	0,32479523	0,0231862		0,94568466239168	4,816085	0,7889	0,092903	1,284662	0,321155
	0,352317	0,18130085	0,1553186		0,5107595532198	5,164245	0,950753	0,863153	1,791254	0,360518
	0,15512	0,29473059	0,1553186	0,074137537	1,03019977867086	3,724188	1,413232	0,944715	2,627219	1,102648
	0,064667	0,15088282	0,1095583	0,168897839	0,41986536763096	3,048974	0,657315	1,839527	2,344618	1,32835
	0,072444	0,29800708	0,0228726	0,123527755	0,3031251086722	1,09004	2,153486	0,410582	1,458271	0,931733
	0,095364		0,0365017	0,046205437	0,35243640744671	3,57031		0,522769	1,321619	1,003405
	0,088406		0,0228389	0,084631653	0,15574405331907	1,087518		0,493059	2,369717	0,824666
	0,049728		0,0906471		0,18086613580611	2,913595		2,168515		0,889165
	0,086359		0,0955387			2,728044		2,143656		
	0,056891		0,11780406			1,606611		1,71219		
	0,080425		0,2285590			4,839735		2,261275		
	0,224639		0,2716558			3,31182		2,421915		
	0,264015		0,2071479			3,195841		1,880202		
	0,205608		0,1573940			2,720649		2,071624		
	0,197615		0,2104329			2,769608		1,992921		
	0,168524		0,21178015			1,293756		2,148743		
			0,2441022					2,078986		
			0,1699271					2,036749		
			0,0898836					1,522682		
			0,1443055					1,5779		
			0,2509242					1,911728		
			0,1903580					1,158811		

	6PG					**Succinate**				
cultivation	aerobic	2.8% O2	1% O2	0.5% O2	0% O2	aerobic	2.8% O2	1% O2	0.5% O2	0% O2
IDs	(19,20,32,33)	(14)	(23,24,30,31)	(21, 22)	(15,16,28,29)	(19,20,32,33)	(14)	(23,24,30,31)	(21, 22)	(15,16,28,29)
O2 (%)	20,9	2,8	1	0,5	0,1	20,9	2,8	1	0,5	0,1
median	1,213	0,512	0,715	0,932	2,146	0,262	0,805	1,421	1,139	7,317
stdev	0,536	0,437	0,473	0,685	0,812	0,091	0,274	0,869	0,660	1,817
stdev%	44	85	66	74	38	35	34	61	58	25
	1,501575	0,25529196	0,1232005	2,759557026	2,71768240737434	0,255502	1,166431	0,534062	0,596273	7,021365
	1,208341	0,15757513	0,3681313	1,038974815	1,7920741277315	0,233801	0,938715	0,401441	2,428058	4,314776
	1,637254	0,23549497	0,1672479	0,647314356	2,35560540768281	0,299314	0,80485	0,086986	1,139483	6,683457
	1,949909	0,64428663	0,78754115	1,10702922	1,79413554778076	0,481039	0,75902	1,776999	1,575803	5,174297
	0,971955	0,88827517	0,7146590	0,93165958	3,89281673680162	0,325229	0,726243	1,810207	1,988831	10,36934
	0,730629	0,51212263	0,2432205	1,142557861	2,53491114722084	0,349172	0,319874	2,204277	1,715067	8,658859
	1,463817	1,38019351	0,3445197	0,671898339	1,93595948158479	0,305676	1,051107	0,44891	0,579698	8,113153
	0,817609		0,2228667	0,387248751	3,13577183869642	0,262977		0,242414	0,690975	9,271077
	1,398155		0,5275757	0,802866737	1,75648484924921	0,237437		0,427991	1,02674	7,597567
	1,22846		0,8929286		1,06505053018325	0,272554		1,631445		7,035564
	2,251238		0,5157394			0,206311		2,071014		
	2,493247		1,0990388			0,228259		2,161861		
	0,822555		0,7662036			0,409829		1,425934		
	1,216864		0,7957949			0,261286		1,416247		
	0,650902		0,7661630			0,354338		1,113557		
	0,939183		0,8297364			0,168878		1,106445		
	0,712678		0,7154821			0,137921		3,452152		
	0,9261		0,7152325			0,121166		2,361766		
			0,4826251					2,602134		
			0,6023225					2,084406		
			1,0549025					2,592576		
			2,2076037					1,137202		
			1,6294552					1,272511		
			1,2332693					1,091897		

	Malate					**Citrate**				
cultivation	aerobic	2.8% O2	1% O2	0.5% O2	0% O2	aerobic	2.8% O2	1% O2	0.5% O2	0% O2
IDs	(19,20,32,33)	(14)	(23,24,30,31)	(21, 22)	(15,16,28,29)	(19,20,32,33)	(14)	(23,24,30,31)	(21, 22)	(15,16,28,29)
O2 (%)	20,9	2,8	1	0,5	0,1	20,9	2,8	1	0,5	0,1
median	0,968	3,401	2,219	1,782	4,608	6,850	13,659	11,036	7,170	13,061
stdev	0,254	1,020	3,743	0,536	1,491	2,601	3,995	5,155	2,664	3,395
stdev%	26	30	169	30	32	38	29	47	37	26
	1,049494	3,87282338	0,17112808	0,7296797	5,77044355686357	4,706798	16,17477	4,297436	7,435344	16,79871
	0,949737	3,40141823	0,4518696	2,001457739	3,46488045881703	11,01382	13,65919	7,021875	13,7689	14,76147
	1,134708	3,38482111	0,1361977	1,123844939	6,3109198903442	10,79916	9,817472	3,257142	7,169836	14,36477
	1,901543	2,38983653	15,052285	1,861740804	2,88217418064318	13,07322	11,79263	22,01639	11,07247	11,27644
	0,968124	3,64346578	14,150215	1,919277287	7,96035928604401	7,737704	19,1668	21,84285	7,384933	20,29975
	0,7998	1,8647571	2,0157627	1,781750598	4,77778969151859	7,085333	7,300603	13,69939	7,151735	11,75632
	1,02224	5,00989948	0,41911806	0,962328214	4,123698185438 93	5,183403	14,8969	6,644675	5,343153	9,925113
	1,006427		0,2273723	1,000161501	4,4389914402569	5,376907		5,124798	5,989437	11,04533
	0,950553		0,4246213	2,119739919	4,79596485999658	5,533378		6,679	7,062823	15,94474
	1,071437		2,0210353		3,92348982177293	5,320009		11,9736		10,08835
	0,968475		1,7933734			4,991861		12,38364		
	1,234841		2,4628037			4,172719		15,0286		
	0,949148		3,4918635			11,1315		11,48115		
	1,218053		3,51122207			6,614405		12,34558		
	0,763406		2,8654972			7,665904		10,59061		
	0,913215		2,8769070			7,246105		9,950038		
	0,938523		3,0384065			6,210321		13,68093		
	0,76426		2,9261458			8,698351		14,04722		
			2,7905233					14,36176		
			2,7436559					13,07376		
			2,1019087					4,430639		
			2,3040131					5,368268		
			2,1332804					7,134258		
			1,3937641					5,155357		

Appendix B

Appendix to Chapter 4

B.1 Appendix to Section 4.1

It seems useful to list the known misprints in the original print article.

Kinetic law for ADH:
(1) the species a should denote NAD and b should denote Ethanol
(2) the last term in the equation should read $bpq/(K \cdot ibK \cdot iqK \cdot p)$

Kinetic law for PFK:
(1)the equation for R should read $R = 1 + \lambda_1 + \lambda_2 + g_r\lambda_1\lambda_2$
(2) the equation L should read $L = L0(..)^2(..)^2(..)^2$ instead of $L = L0(..)^2(..)^2(..)$
A further small deviation from the printed article (perhaps misprint) is the steady state concentration of $G6P$ being 1.03 mmol/l rather than 1.07 mmol/l as stated in the paper.

B.2 Appendix to Section 4.2

	G6P	F6P	Trio	PEP	PYR	AcA	OGA	S7P	E4P	acetate	$AcCoA_{mit}$	$AcCoA_{cyt}$	Oaa_{mit}	Oaa_{cyt}	cytrate	CO_2	ATP	NADH	O_2	R5P
v1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0
v2	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
v3	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
v4	0	-1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0
v5	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	-2
v6	0	1	1	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	-1
v7	0	1	-1	0	0	0	0	-1	1	0	0	0	0	0	0	0	0	0	0	0
v8	0	0	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
v9	0	0	0	-1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
v10	0	0	0	0	-1	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0
v11	0	0	0	0	0	0	0	0	0	0	-1	0	-1	0	1	0	0	0	0	0
v12	0	0	0	0	0	0	1	0	0	0	0	0	0	0	-1	1	0	1	0	0
v13	0	0	0	0	0	0	-1	0	0	0	0	0	1	0	0	1	$\frac{2}{3}R_{p/o}$	2	$-\frac{1}{2}$	0
v14	0	0	0	0	1	0	0	0	0	0	0	0	-1	0	0	1	0	0	0	0
v15	0	0	0	1	0	0	0	0	0	0	0	0	0	-1	0	1	-1	0	0	0
v16	0	0	0	0	-1	0	0	0	0	0	0	0	0	1	0	-1	-1	0	0	0
v17	0	0	0	0	0	0	0	0	0	-1	0	1	0	0	0	0	-1	0	0	0
v18	0	0	0	0	0	-1	0	0	0	1	0	0	0	0	0	0	0	1	0	0
v19	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	-1	0	0
v20	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0
v21	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0
v22	0	0	0	0	0	0	0	0	0	0	0	0	-1	1	0	0	0	0	0	0
v24	0	0	0	0	-1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
v29	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0
v31	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0
v30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1
v32	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
v33	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
v34	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0
v35	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0
v38	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0
v40	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
v51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
v52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0
v53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$R_{p/o}$	-1	$-\frac{1}{2}$	0
v54	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0
	G6P	F6P	Trio	PEP	PYR	AcA	OGA	S7P	E4P	acetate	$AcCoA_{mit}$	$AcCoA_{cyt}$	Oaa_{mit}	Oaa_{cyt}	cytrate	CO_2	ATP	NADH	O_2	R5P

Table B.1: The transposed of the stoichiometrix matrix used in the constrained flux balance procedure. Tables 4.1(l) and 4.1(r), and Figure 4.4 show other representations of this network. $R_{p/o}$ denotes the p/o-ratio which was set to unity. See also Footnote 3 in Chapter 4.

reaction	0%	0.5 %	1 %	2.8 %	20.9%
v1	68.2500	23.8333	21.3417	14.7333	13.8667
v2	60.1585	17.5965	14.5495	7.2058	7.3778
v3	3.3140	3.3768	3.1641	3.9915	3.1609
v4	60.0914	18.5886	14.1456	7.6823	7.2634
v5	1.3318	0.6791	0.9579	0.7294	0.9331
v6	0.6493	0.4408	0.5311	0.2874	0.6558
v7	1.3318	0.6791	0.9579	0.7294	0.9331
v8	107.1228	36.6644	27.7548	14.7679	14.3506
v9	106.4997	37.8692	27.1150	14.1593	14.2350
v10	1.8101	4.2537	5.9956	8.8800	9.6346
v11	1.8101	4.2537	5.9956	8.8800	9.6346
v12	1.8101	4.2537	5.9956	8.8800	9.6346
v13	-0.0000	2.8253	4.2943	7.1151	7.9698
v14	0.0000	-0.0000	0.0000	0.5176	0.4776
v15	-0.0000	1.9197	-0.0000	0.2752	0.7164
v16	3.1158	4.5392	3.1931	4.0294	4.1072
v17	1.2463	0.7147	1.0663	1.0309	0.9707
v18	1.2463	0.7147	1.0663	1.0309	0.9707
v19	100.3275	28.3616	16.8599	0.7366	0.0001
v20	11.6618	0.0001	0.0004	0.0002	0.0000
v21	1.8101	4.3344	3.7292	8.2890	7.7241
v22	-0.0000	2.9060	2.0278	6.0065	5.5817
v24	101.5738	29.0763	17.9262	1.7675	0.9708
v29	4.7775	2.8600	3.6281	3.5360	3.3280
v30	0.0009	1.5779	0.7172	2.2454	0.6388
v31	0.6825	0.2383	0.4268	0.4420	0.2773
v32	2.0475	0.9533	1.0671	0.8840	0.8320
v33	0.6232	0.7148	0.6399	0.8838	0.8320
v34	1.3057	1.1911	1.4918	1.4717	1.2484
v35	1.2463	0.7147	1.0663	1.0309	0.9707
v38	1.8101	1.4284	1.7013	1.7649	1.6648
v40	2.0483	0.1278	1.8929	0.5402	1.7033
v51	-0.0000	13.0003	18.4175	27.0836	29.2499
v52	105.3922	41.1663	34.1827	27.3974	28.4575
v53	0.0000	23.1753	32.5406	47.0522	50.5301
v54	76.1415	47.1370	46.8985	49.4356	54.1764

Table B.2: Numerical values given in [mM/min] for the result of the constrained flux balance analysis procedure described in Section 4.2.3 for each oxygenation condition, as presented in Fig. 4.5. The objective function Eq. 4.12 is evaluated to 7.7355 10^{-5}. These flux distributions are steady state solutions for the stoichiometric network represented by Table B.1 with residual fluxes lower than O(10^{-13}).

The kinetic equations of the model (based on [87]). Variables are denoted by *slanted* font,
parameters are denoted by `typewriter` font. Variables denote concentrations in [mM]; reaction rates are in [mM/min].

$$vGLK = \mathtt{VmaxGLK} \cdot \frac{(GLCiATP - G6PADP/\mathtt{KeqGLK})}{(\mathtt{KmGLKGLCi} \cdot \mathtt{KmGLKATP})} \cdot \frac{1}{(1 + GLCi/\mathtt{KmGLKGLCi} + G6P/\mathtt{KmGLKG6P})} \cdot \frac{1}{(1 + ATP/\mathtt{KmGLKATP} + ADP/\mathtt{KmGLKADP})} \tag{B.1}$$

$$vPGI = \mathtt{VmaxPGI}/\mathtt{KmPGIG6P}\,(G6P - F6P/\mathtt{KeqPGI}) \Big/ (1 + G6P/\mathtt{KmPGIG6P} + F6P/\mathtt{KmPGIF6P})$$

$$vPFK = \mathtt{VmPFK}\ \mathtt{g_R}\,(F6P/\mathtt{KmPFKF6P})\,(ATP/\mathtt{KmPFKATP})\,R_{PFK}/(R_{PFK}^2 + L_{PFK}\,T_{PFK}^2)$$

where $R_{PFK} := 1 + \dfrac{F6P}{\mathtt{K_{R,F6P}}} + \dfrac{ATP}{\mathtt{K_{R,ATP}}} + \mathtt{g_R}\dfrac{F6P}{\mathtt{K_{R,F6P}}}\dfrac{ATP}{\mathtt{K_{R,ATP}}}$ $\qquad T_{PFK} := 1 + \mathtt{C}_{i\mathtt{ATP}}\dfrac{ATP}{\mathtt{K}_{R,\mathtt{ATP}}}$

$$L_{PFK} := \mathtt{L_0}\left(\frac{1 + \mathtt{C}_{i\mathtt{ATP}}ATP/\mathtt{K_{ATP}}}{1 + ATP/\mathtt{K_{ATP}}}\right)^2 \cdot \left(\frac{1 + \mathtt{C}_{i\mathtt{AMP}}AMP/\mathtt{K_{AMP}}}{1 + AMP/\mathtt{K_{AMP}}}\right)^2 \cdot \left(\frac{1 + \mathtt{C}_{i\mathtt{F26bP}}F26bP/\mathtt{K_{F26bP}} + \mathtt{C_{F16bP}}F16bP/\mathtt{K_{F16bP}}}{1 + F26bP/\mathtt{K_{F26bP}} + F16bP/\mathtt{K_{F16bP}}}\right)^2$$

$$vALD = \frac{\mathtt{VmaxALD}}{\mathtt{KmALDF16P}} \cdot \frac{(F16P - GAP\ DHAP/\mathtt{KeqALD})}{1 + \frac{F16P}{\mathtt{KmALDF16P}} + \frac{GAP}{\mathtt{KmALDGAP}} + \frac{DHAP}{\mathtt{KmALDDHAP}} + \frac{GAP\ DHAP}{\mathtt{KmALDGAP\ KmALDDHAP}} + \frac{F16P\ GAP}{\mathtt{KmALDGAP}i\ \mathtt{KmALDF16P}}}$$

$$vG3PDH = \frac{\mathtt{VmG3PDH}}{(\mathtt{KmG3PDHDHAP} \cdot \mathtt{KmG3PDHNADH})} \frac{(DHAP \cdot NADH - GLY\ NAD/\mathtt{KeqG3PDH})}{(1 + DHAP/\mathtt{KmG3PDHDHAP} + GLY/\mathtt{KmG3PDHGLY}) \cdot (1 + NADH/\mathtt{KmG3PDHNADH} + NAD/\mathtt{KmG3PDHNAD})}$$

with the concentrations GAP and $DHAP$ calculated from the equilibrium assumption of TPI:
$DHAP := TRIO/(1 + \mathtt{KeqTPI})$ and $GAP := \mathtt{KeqTPI}\,TRIO/(1 + \mathtt{KeqTPI})$ **(continued on next page)**

(continued from last page)

$$vGAPDH = \frac{\texttt{VmGAPDHf}}{\texttt{KmGAPDHGAP KmGAPDHNAD}} GAP\ NAD - \texttt{VmGAPDHr}\ BPG\ NADH$$

with the rewerse maximal velocity `VmGAPDHr` calculated via the Haldane-relationship:

$$\texttt{VmGAPDHr} := \frac{\texttt{VmGAPDHf}}{\texttt{vGAPDH\_Keq}} (\texttt{KmGAPDHBPG KmGAPDHNADH}) / (\texttt{KmGAPDHGAP KmGAPDHNAD})$$

$$vPGK = \frac{\texttt{VmPGK}}{KmPGKP3G\ \texttt{KmPGKATP}} \cdot \frac{\texttt{KeqPGK}\ BPG\ ADP - P3G\ ATP}{(1 + BPG/\texttt{KmPGKBPG} + P3G/\texttt{KmPGKP3G})(1 + ATP/\texttt{KmPGKATP} + ADP/\texttt{KmPGKADP})}$$

$$vPGM = \frac{\texttt{VmPGM}}{\texttt{KmPGMP3G}} \cdot \frac{P3G - P2G/\texttt{KeqPGM}}{1 + P3G/\texttt{KmPGMP3G} + P2G/\texttt{KmPGMP2G}}$$

$$vENO = \frac{\texttt{VmENO}}{\texttt{KmENOP2G}} \cdot \frac{P2G - PEP/\texttt{KeqENO}}{1 + P2G/\texttt{KmENOP2G} + PEP/\texttt{KmENOPEP}}$$

$$vPYK = \frac{\texttt{VmPYK}}{\texttt{KmPYKPEP KmPYKADP}} \cdot \frac{PEP\ ADP - PYR\ \ ATP/\texttt{KeqPYK}}{(1 + PEP/\texttt{KmPYKPEP} + PYR/\texttt{KmPYKPYR}) * (1 + ATP/\texttt{KmPYKATP} + ADP/\texttt{KmPYKADP})}$$

$$vTCA = \frac{\texttt{vTCA\_Vmax}}{(1 + \frac{TCAint}{\texttt{vTCA\_KiTCAint}})} \cdot \frac{1}{(1 + \frac{\texttt{vTCA\_KmPyr}}{Pyr}) \cdot (1 + \frac{\texttt{vTCA\_KmNAD}}{NAD})}$$

$$vRESP = \texttt{vRESP\_K} \cdot O_2 \cdot NADH \cdot ADP$$

$$vRESP2 = \texttt{vRESP2\_K} \cdot O_2 \cdot TCAint \cdot ADP \cdot NAD$$

$$vATPase = \texttt{vATPase\_K} \cdot ATP$$

(continued on next page)

The parameter values are as follows:

adenosin kinase (in equilibrium): KeqAK = 0.45
triosephosphate isomerase (in equilibrium): KeqTPI = 0.045
vPFK: gR = 5.12 KmPFKF6P = 0.1 KmPFKATP = 0.7 Lzero = 0.66 CiPFKATP = 100 KiPFKATP = 0.65 CPFKAMP = 0.0845 KPFKAMP = 0.0995 CPFKF26BP = 0.0174 KPFKF26BP = 0.000682 CPFKF16BP = 0.397 KPFKF16BP = 0.111 CPFKATP = 3
vGLK: KmGLKGLCi = 0.08 KmGLKATP = 0.15 KeqGLK = 3800 KmGLKG6P = 30 KmGLKADP = 0.23
vPGI: KmPGIG6P = 1.4 KeqPGI = 0.314 KmPGIF6P = 0.3
vALD: KmALDF16P = 0.3 KeqALD = 0.07 KmALDGAP = 2 KmALDDHAP = 2.4 KmALDGAPi = 10
vGAPDH: KmGAPDHGAP = 0.21 KmGAPDHNAD = 0.09 KmGAPDHBPG = 0.0098 KmGAPDHNADH = 0.06 Keq = 0.0056
vPGK: KmPGKP3G = 0.53 KmPGKATP = 0.3 KeqPGK = 3200 KmPGKBPG = 0.003 KmPGKADP = 0.2
vPGM: KmPGMP3G = 1.2 KeqPGM = 0.19 KmPGMP2G = 0.08
vENO: KmENOP2G = 0.04 KeqENO = 6.7 KmENOPEP = 0.5
vPYK: KmPYKPEP = 0.14 KmPYKADP = 0.53 KeqPYK = 6500 KmPYKPYR = 21 KmPYKATP = 1.5
vPDV: nPDC = 2 KmPDCPYR = 4.33
vADH: KiADHNAD = 0.92 KmADHETOH = 17 KeqADH = 6.9 e-05 KmADHNAD = 0.17 KmADHNADH = 0.11 KiADHNADH = 0.031 KmADHACE = 1.11 KiADHACE = 1.1 KiADHETOH = 90
vG3PDH: KmG3PDHDHAP = 0.4 KmG3PDHNADH = 0.023 KeqG3PDH = 4300 KmG3PDHGLY = 1 KmG3PDHNAD = 0.93

species	anaerobic (0% oxygen) nom.value ; [min., max.]	reference,remark	aerobic (20.9% oxygen) nom.value; [min., max.]	reference,remark
(GAP)	0.15	[87]	0.063	[90]
(DHAP)	0.81	[87]	0.34	assuming anaerobic $\frac{[GAP]}{[DHAP]}$ ratio
TRIO (GAP+DHAP)	0.96; $[\frac{1}{2},\ 2]$		0.4, $[\frac{2}{3};\ \frac{3}{2}]$	
BPG	$5.83 \cdot 10^{-4}$; $[\frac{1}{3},\ 3]$	[87], inferred from kinetics	$5.83 \cdot 10^{-4}$; $[\frac{1}{3},\ 3]$	using anaerobic value
(G3P+G2P)	2.07	[97]	4.94	
(G3P:G2P)	9:1	[87]	9:1	assuming anaerobic value
G3P	1.86; $[\frac{2}{3},\ [\frac{3}{2}]$		4.45; $[\frac{2}{5},\ \frac{5}{2}]$	
G2P	0.20; $[\frac{2}{3},\ [\frac{3}{2}]$		0.49; $[\frac{2}{5},\ \frac{5}{2}]$	
ACA	0.17; $[\frac{1}{2},\ 2]$	[87]	0.17; $[\frac{1}{3},\ 3]$	[87]

Table B.3: Objective nominal values and ranges for intermediate metabolite species not measured in [97], as used in the parameter estimation for the kinetic model. Experimental quantities used to infer these are listed in brackets above the corresponding metabolite. Nominal values are given in [mM]; minimal and maximal values are given in multiples of the nominal value. The data are based on the following cell cultures:
[97, Wiebe *et al.*], mostly used in this work: anaerobic and aerobic chemostat culture.
[87, Teusink *et al.*]: anaerobic batch culture.
[90, Theobald *et al.*]: aerobic chemostat culture.

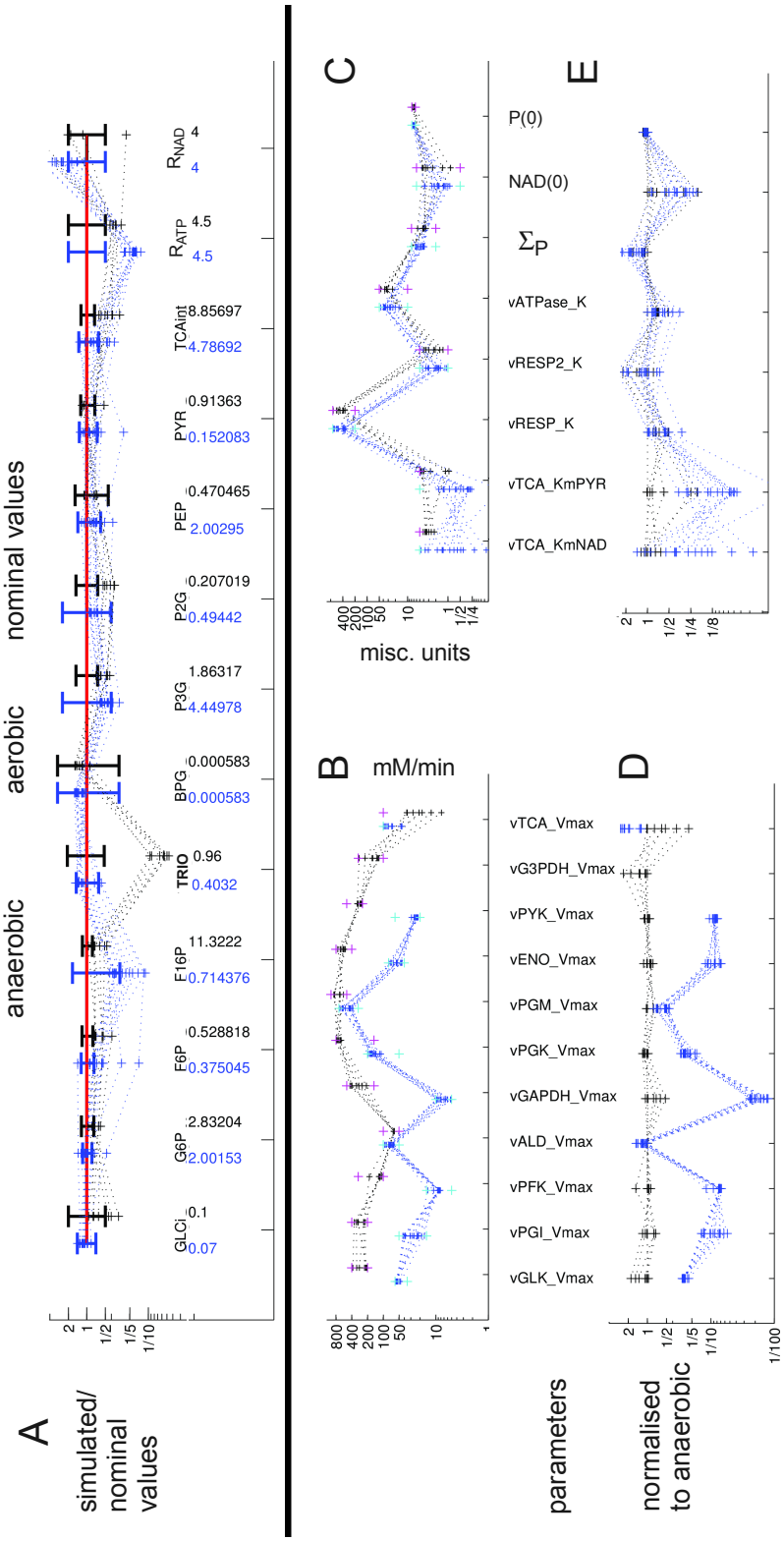

Figure B.1: Result of fitting the kinetic model using experimental datasets from anaerobic (0% oxygen, black) and aerobic (20.9 % oxygen, blue) chemostat yeast cultures showing the corresponding parameter sets. To reduce the volume of parameter space, the reactions *vG3PDH*, *vPDC*, *vADH*, and *vPDbp* were fixed to their nominal values and the corresponding Vmax not included in the search. Shown are results from those 22 parameter searches which had an objective function value below 0.1. The parameter search used the differential evolution algorithm as implemented in `SBTOOLBOX2` for 1000 generations with populations of 3000 parameter sets.
Panel A: Fitted quantities: Simulated values plotted on a logarithmic scale as multiples of corresponding nominal values (appearing as unity, red line). Names and absolute nominal values (colour coded) are indicated. R_{ATP} and R_{NAD} denote ATP/ADP and NAD/NADH ratio, respectively. Uncertainties of nominal values are denoted by error bars. Error bars and simulated values are colour coded to indicate oxygenation conditions.
Panels B, C, D, E: Estimated Parameters. Panels B and C show absolute parameter values plotted on a logarithmic scale. Upper and lower limits of the search range for each parameter are shown as light blue (aerobic) and magenta (anaerobic) crosses. V_{max} values (Panel B) are shown in mM/min, dimensions for the rest of the parameters (Panel C) vary.
Panels D, E show parameter values plotted on a logarithmic scale normalised to the corresponding value from one arbitrary chosen anaerobic parameter set.

species or reaction	anaerobic value	anaerobic value
Ethanol	75.3708	0
Glycerol	8.9025	0
O2	0	0.25
vGLT	68.25	13.8667
v29	4.7775	3.328
v3	3.3129	3.6254
v5	1.3318	0.9331
v6	0.6493	0.6558
v7	1.3318	0.9331
v15	0	0.7158
v16	3.1158	4.1067
v40	2.0492	1.2387
v32	2.0475	0.832
v33	0.6232	0.832
v38	1.8101	1.6648

Table B.4: External species and reactions and their set values (shown in [mM] and [mM/min], respectively) for modelling the anaerobic and the aerobic scenario, c.f. Section 4.2.5. For calculating mM from μmol/(g dry weight) in the publications, see Section 2.7.

B.3 Appendix to Section 4.3

$$
\begin{aligned}
HK =& (HXK1 + HXK2 + GLK1)/(HXK1_0 + HXK2_0 + GLK1_0)\\
PGI =& normdPGI1\\
PFK =& (PFK1 + PFK2)/(PFK1_0 + PFK2_0)\\
FBP =& normdFBP1\\
ALD =& normdFBA1\\
GlyceF =& (GPD1 + GPD2 + RHR2 + HOR2)/(GPD1_0 + GPD2_0 + RHR2_0 + HOR2_0)\\
GlyceR =& normdGUT1\\
GAPDH =& (TDH1 + TDH3)/(TDH1_0 + TDH3_0)\\
PGK =& normdPGK1\\
PGM =& normdGPM1\\
ENO =& (ENO1 + ENO2)/(ENO1_0 + ENO2_0)\\
PYK =& (CDC19 + PYK2)/(CDC19_0 + PYK2_0)\\
PDC =& (PDC1 + PDC5 + PDC6 + ARO10 + THI3)/(PDC1_0 + PDC5_0 + PDC6_0 + ARO10_0 + THI3_0)\\
PYC =& (PYC1 + PYC2)/(PYC1_0 + PYC2_0)\\
ADH =& \frac{(ADH1 + ADH3 + ADH4 + ADH5 + ADH6 + ADH7)}{(ADH1_0 + ADH3_0 + ADH4_0 + ADH5_0 + ADH6_0 + ADH7_0)}\\
PDHc =& YIA6 + \frac{(PDA1 + PDB1 + LPD1 + LAT1)}{(YIA6_0 + PDA1_0 + PDB1_0 + LPD1_0 + LAT1_0)}\\
TCA1 =& 1/3 \cdot (normdYIA6 + \frac{(PDA1 + PDB1 + LPD1 + LAT1 + PDX1)}{(PDA1_0 + PDB1_0 + LPD1_0 + LAT1_0 + PDX1_0)}\\
&+ (CIT1 + CIT3)/(CIT1_0 + CIT3_0) + (ACO1 + ACO2)/(ACO1_0 + ACO2_0)\\
&+ \frac{(IDH1 + IDH2 + IDP1 + IDP2 + IDP3)}{(IDH1_0 + IDH2_0 + IDP1_0 + IDP2_0 + IDP3_0)})
\end{aligned}
\tag{B.2}
$$

Appendix C

Abbreviations

Enzymes:

Abbrev.	(alternative abbreviations) name, explanation
ADH	alcohol dehydrogenase
ALD	aldolase
ENO	enolase
FBP	fructose bisphosphatase
Fum	fumarase
GAPDH	glyceraldahyde-3-P dehydrogenase
G3PDH	glycerol-3-phosphate dehydrogenase
HK	hexokinase
PCK	phosphoenolpyruvate carboxykinase
PDC	pyruvate decarboxylase
PDHc	(PDH-complex) pyruvate dehydrogenase complex
PFK	phosphofructokinase
PGI	phosphoglucoisomerase
PGK	phosphoglycerate kinase
PGM	phosphoglycerate mutase
PYC	pyruvate carboxylase
PYK	pyruvate kinase
PDbp	pyruvate dehydrogenase bypass

Metabolites:

AcA	acetaldehyde
Acetyl-Coa	acetyl coenzyme A
ATP/ ADP/ AMP	Adenosine-tri/di/monophosphate
cAMP	cyclic AMP
CIT	citrate
CO_2	carbon-dioxide
DHAP	dihydroxyacetone phosphate
E4P	Erythrose-4-phosphate
F16bP	(F1,6bP) fructose-1,6-bisphosphate
F26bP	(F2,6bP) fructose-2,6-bisphosphate
F6P	fructose 6-phosphate
Fum	fumarate
G6P	glucose-6-phosphate
GAPD	glyceraldehyde-3-phosphate
ICIT	isocitric acid
Mal	malate
NAD, NADH	(NAD^+) nicotinamide adenine dinucleotide
O_2	oxygen (molecular)
Oaa	oxaloacetic acid
OGA	(AKG) α-ketoglutarate
PEP	phosphoenolpyruvate
Pyr	pyruvate
R5P	ribulose-5-phosphate
S7P	sedoheptulose 7-phosphate
Trio	GAP+DHAP (see Chpt. 4)

other:

TCA	cytic acid cycle
PPP	pentose phosphate pathway
Cys	cysteine

Acknowledgement

First and foremost I wish to thank my family who supported me throughout the time leading to this work as well as Cecilia (Marta Hardler Bertoni) Chirenti.

I wish to acknowledge the support and the influence on my work of the late Prof. Reinhardt Heinrich during the initial period of my PhD, as well as of my supervisor Prof. Eda Klipp during the subsequent years who proposed the topic of this thesis and initiated cooperation with the experimental partner. I also thank her for the many extended discussions we had on the subject. I also wish to thank Dr. Juha-Pekka Pitänen for taking the time for discussions, hence granting me access to his extensive knowledge, and hosting me during my visit to Espoo/Helsinki, as well as to Paula Jouhten, Dr. Marilyn Wiebe, and Prof. Merja Penttilä of VTT Research Centre of Finland for insightful discussions on topics related to yeast cultivation.

While a number of other people in the group (Lehrstuhl) of Theoretische Biophysik supported me in directly or indirectly, I especially would like to thank Dr. Bernd Binder for critically reviewing most of the final manuscript, Dr. Carl-Fredrik Tiger for critically reviewing Chapter two, as well as Dr. Clemens Kühn, Dr. Matteo Barberis, Dr. Axel Kowald, Dr. François Guillaud, Margot Heiske, and Dr. Samuel Drulhe for supporting me with both their expertise and providing a cheerful work environment (and Axel even his desk!).

I wish to thank my friends outside my professional circles for invaluable support, such as Jan-Yves Weseloh, Susi Szymkowiak, Denny Milakara, Gülşen Fidan, and Olli Sabiniarz, but I also wish to thank to Mark Heinzle, Natascha Rahi, Christian Lahl in Vienna, as well as to Zsófi József, Robi Vértesi and other friends in Budapest for not having abandoned me during all my years abroad (respectively).

Financial support (June. 2005 - Dec. 2008) granted under Prof. R. Heinrich's supervision by the Marie Curie Early Stage Training Project: Systems Biology (contract number MEST-2-CT-2004-514169), financial support (Jan. - June 2010) granted under Prof. E. Klipp's supervision by the Marie Curie Early Stage Training Project Systems Biology; UNICELLSYS (Contract No. 201142), and financial support (Oct. - Nov. 2010) by the Academy of Finland and TEKES (project number 1050/31/07, Biocontrol) is gratefully acknowledged.

The author wishes to gratefully acknowledge the contribution of the European, especially German and Finnish taxpayer in making this work possible.

Bibliography

[1] *SGD - Saccharomyces Genome Database*. http://www.yeastgenome.org/

[2] Alberts, B. ; Johnson, A. ; Lewis, J. ; Raff, M. ; Roberts, K. ; Walter, P.: *Molecular Biology of the Cell*. 4th. New York : Garland Science, 2002. – ca. 1500 S.

[3] Aragón, J J. ; Gómez, M E. ; Gancedo, C: Identification of two forms of 6-phosphofructo-2-kinase in yeast. In: *FEBS letters* 226 (1987), Dezember, Nr. 1, 121–4. http://www.ncbi.nlm.nih.gov/pubmed/2961616. – ISSN 0014–5793

[4] Bartel, David P.: MicroRNAs: target recognition and regulatory functions. In: *Cell* 136 (2009), Januar, Nr. 2, 215–33. http://dx.doi.org/10.1016/j.cell.2009.01.002. – DOI 10.1016/j.cell.2009.01.002. – ISSN 1097–4172

[5] Beyer, Andreas ; Hollunder, Jens ; Nasheuer, Heinz-Peter ; Wilhelm, Thomas: Post-transcriptional expression regulation in the yeast Saccharomyces cerevisiae on a genomic scale. In: *Molecular & cellular proteomics : MCP* 3 (2004), November, Nr. 11, 1083–92. http://dx.doi.org/10.1074/mcp.M400099-MCP200. – DOI 10.1074/mcp.M400099–MCP200. – ISSN 1535–9476

[6] Blank, Lars M. ; Kuepfer, Lars ; Sauer, Uwe: Large-scale 13C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. In: *Genome biology* 6 (2005), Januar, Nr. 6, R49. http://dx.doi.org/10.1186/gb-2005-6-6-r49. – DOI 10.1186/gb–2005–6–6–r49. – ISSN 1465–6914

[7] Bordel, Sergio ; Agren, Rasmus ; Nielsen, Jens: Sampling the solution space in genome-scale metabolic networks reveals transcriptional regulation in key enzymes. In: *PLoS computational biology* 6 (2010), Januar, Nr. 7, e1000859. http://dx.doi.org/10.1371/journal.pcbi.1000859. – DOI 10.1371/journal.pcbi.1000859. – ISSN 1553–7358

[8] Borger, Simon: *Parametrising kinetic models of biological networks*, Humboldt Universität zu Berlin, Diss., 2008. http://edoc.hu-berlin.de/docviews/abstract.php?id=30349

[9] Brink, Joost Van D. ; Canelas, B ; Gulik, Walter M V. ; Pronk, Jack T. ; Heijnen, Joseph J. ; Winde, Johannes H D. ; Daran-lapujade, Pascale: Dynamics of Glycolytic Regulation during Adaptation of Saccharomyces cerevisiae to Fermentative Metabolism â . In: *Applied and environmental microbiology* 74 (2008), Nr. 18, S. 5710–5723. http://dx.doi.org/10.1128/AEM.01121-08. – DOI 10.1128/AEM.01121–08

[10] Canelas, André B ; Gulik, Walter M. ; Heijnen, Joseph J.: Determination of the cytosolic free NAD/NADH ratio in Saccharomyces cerevisiae under steady-state and highly dynamic conditions. In: *Biotechnology and bioengineering* 100 (2008), Juli, Nr. 4, 734–43. http://dx.doi.org/10.1002/bit.21813. – DOI 10.1002/bit.21813. – ISSN 1097–0290

[11] Cerdan, Sebastian ; Gancedo, Juana M.: Futile cycles in Saccharomyces cerevisiae strains expressing the gluconeogenic enzymes during growth on glucose. In: *Biochemistry* 90 (1993), Nr. February, S. 1290–1294

[12] Chassagnole, Christophe ; Noisommit-Rizzi, Naruemol ; Schmid, Joachim W. ; Mauch, Klaus ; Reuss, Matthias: Dynamic modeling of the central carbon metabolism ofEscherichia coli. In: *Biotechnology and Bioengineering* 79 (2002), Juli, Nr. 1, 53–73. http://dx.doi.org/10.1002/bit.10288. – DOI 10.1002/bit.10288. – ISSN 0006–3592

[13] Cornish-Bowden, Athel: *Fundamentals of Enzyme Kinetics*. Portland Press. – 438 S. – ISBN 978–1855781580

[14] Crécy, E de ; Metzgar, D ; Allen, C ; Pénicaud, M ; Lyons, B ; Hansen, C J. ; Crécy-Lagard, V de: Development of a novel continuous culture device for experimental evolution of bacterial populations. In: *Applied microbiology and biotechnology* 77 (2007), November, Nr. 2, 489–96. http://dx.doi.org/10.1007/s00253-007-1168-5. – DOI 10.1007/s00253–007–1168–5. – ISSN 0175–7598

[15] Dalma-Weiszhausz, Dennise D. ; Warrington, Janet ; Tanimoto, Eugene Y. ; Miyada, C G.: The affymetrix GeneChip platform: an overview. In: *Methods in enzymology* 410 (2006), Januar, 3–28. http://dx.doi.org/10.1016/S0076-6879(06)10001-4. – DOI 10.1016/S0076–6879(06)10001–4. – ISSN 0076–6879

[16] Dam, J van ; Eman, Michael R. ; Frank, Johannes ; Lange, Hans C. ; Dedem, Gijs W. K. ; Heijnen, Sef J.: Analysis of glycolytic intermediates in Saccharomyces cerevisiae using anion exchange chromatography and electrospray ionization with tandem mass spectrometric detection. In: *Analytica Chimica Acta* 460 (2002), Juni, Nr. 2, 209–218. http://dx.doi.org/10.1016/S0003-2670(02)00240-4. – DOI 10.1016/S0003–2670(02)00240–4. – ISSN 00032670

[17] Danø, S ; Sø rensen, P G. ; Hynne, F: Sustained oscillations in living cells. In: *Nature* 402 (1999), Nr. 6759, 320–322. http://www.ncbi.nlm.nih.gov/pubmed/10580506

[18] Daran-lapujade, P ; Daran, J ; Maris, A van ; Winde, J de ; Pronk, J: Chemostat-Based Micro-Array Analysis in Baker's Yeast. In: *Advances in Microbial Physiology* 54 (2008), Nr. 08, 257–417. http://dx.doi.org/10.1016/S0065-2911(08)00004-0. – DOI 10.1016/S0065–2911(08)00004–0. – ISSN 00652911

[19] De Koning, W ; Van Dam, K: A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH. In: *Analytical Biochemistry* 204 (1992), Nr. 1, 118–123. http://www.ncbi.nlm.nih.gov/pubmed/1514678

[20] Dihazi, H ; Kessler, R ; Eschrich, K: Phosphorylation and inactivation of yeast 6-phosphofructo-2-kinase contribute to the regulation of glycolysis under hypotonic stress. In: *Biochemistry* 40 (2001), Dezember, Nr. 48, 14669–78. http://www.ncbi.nlm.nih.gov/pubmed/11724581. – ISSN 0006–2960

[21] Dihazi, Hassan ; Kessler, Renate ; Eschrich, Klaus: High osmolarity glycerol (HOG) pathway-induced phosphorylation and activation of 6-phosphofructo-2-kinase are essential for glycerol accumulation and yeast cell proliferation under hyperosmotic stress. In: *The Journal of biological chemistry* 279 (2004), Juni, Nr. 23, 23961–8. http://dx.doi.org/10.1074/jbc.M312974200. – DOI 10.1074/jbc.M312974200. – ISSN 0021–9258

[22] Duve, Christian de: *Die Zelle: Expedition in die Grundstruktur des Lebens*. Spektrum Akademischer Verlag, 1993. – 455 S. – ISBN 978–3860250716

[23] Einstein, Albert ; Born, Max ; Born, Hedwig: *Letter from Einstein to Max Born, 3 March 1947, in: The Born-Einstein letters; correspondence between Albert Einstein and Max and Hedwig Born from 1916 to 1955.* New York, 1971

[24] Érdi, Péter ; Tóth, János: *Mathematical Models of Chemical Reactions. Theory and Applications of Deterministic and Stochastic Models*. Princeton : Princeton University Press, 1989

[25] Eunen, Karen van ; Bouwman, Jildau ; Daran-Lapujade, Pascale ; Postmus, Jarne ; Canelas, André B ; Mensonides, Femke I C. ; Orij, Rick ; Tuzun, Isil ; Brink, Joost van d. ; Smits, Gertien J. ; Gulik, Walter M. ; Brul, Stanley ; Heijnen, Joseph J. ; Winde, Johannes H. ; Mattos, M Joost T. ; Kettner, Carsten ; Nielsen, Jens ; Westerhoff, Hans V. ; Bakker, Barbara M.: Measuring enzyme activities under standardized in vivo-like conditions for systems biology. In: *The FEBS journal* 277 (2010), Februar, Nr. 3, 749–60. http://dx.doi.org/10.1111/j.1742-4658.2009.07524.x. – DOI 10.1111/j.1742–4658.2009.07524.x. – ISSN 1742–4658

[26] Fell, David (Oxford Brookes U.: *Understanding the Control of Metabolism*. 1996. – 300 S. – ISBN 185578047X

[27] Fournier, Marjorie L. ; Paulson, Ariel ; Pavelka, Norman ; Mosley, Amber L. ; Gaudenz, Karin ; Bradford, William D. ; Glynn, Earl ; Li, Hua ; Sardiu, Mihaela E. ; Fleharty, Brian ; Seidel, Christopher ; Florens, Laurence ; Washburn, Michael P.: Delayed correlation of mRNA and protein expression in rapamycin-treated cells and a role

for Ggc1 in cellular sensitivity to rapamycin. In: *Molecular & cellular proteomics : MCP* 9 (2010), Februar, Nr. 2, 271–84. http://dx.doi.org/10.1074/mcp.M900415-MCP200. – DOI 10.1074/mcp.M900415–MCP200. – ISSN 1535–9484

[28] François, J ; Van Schaftigen, E ; Hers, H G.: Characterization of phosphofructokinase 2 and of enzymes involved in the degradation of fructose 2,6-bisphosphate in yeast. In: *European journal of biochemistry / FEBS* 171 (1988), Februar, Nr. 3, 599–608. http://www.ncbi.nlm.nih.gov/pubmed/2831055. – ISSN 0014–2956

[29] Gancedo, C ; Flores, C: The importance of a functional trehalose biosynthetic pathway for the life of yeasts and fungi. In: *FEMS Yeast Research* 4 (2004), Januar, Nr. 4-5, 351–359. http://dx.doi.org/10.1016/S1567-1356(03)00222-8. – DOI 10.1016/S1567–1356(03)00222–8. – ISSN 15671356

[30] Gánti, Tibor: Organisation of chemical reactions into dividing and metabolizing units: the chemotons. In: *Biosystems* 7 (1975), S. 189 – 95

[31] Gerosa, Luca ; Sauer, Uwe: Regulation and control of metabolic fluxes in microbes. In: *Current opinion in biotechnology* 22 (2011), Mai, Nr. 4, 566–575. http://dx.doi.org/10.1016/j.copbio.2011.04.016. – DOI 10.1016/j.copbio.2011.04.016. – ISSN 1879–0429

[32] Gombert, A K. ; Moreira dos Santos, M ; Christensen, B ; Nielsen, J: Network identification and flux quantification in the central metabolism of Saccharomyces cerevisiae under different conditions of glucose repression. In: *Journal of bacteriology* 183 (2001), Februar, Nr. 4, 1441–51. http://dx.doi.org/10.1128/JB.183.4.1441-1451.2001. – DOI 10.1128/JB.183.4.1441–1451.2001. – ISSN 0021–9193

[33] Gonzalez, B ; François, J ; Renaud, M: A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol. In: *Yeast (Chichester, England)* 13 (1997), November, Nr. 14, 1347–55. http://dx.doi.org/10.1002/(SICI)1097-0061(199711)13:14<1347::AID-YEA176>3.0.CO;2-O. – DOI 10.1002/(SICI)1097–0061(199711)13:14<1347::AID–YEA176>3.0.CO;2–O. – ISSN 0749–503X

[34] Hans, M. a. ; Heinzle, E. ; Wittmann, C.: Quantification of intracellular amino acids in batch cultures of Saccharomyces cerevisiae. In: *Applied Microbiology and Biotechnology* 56 (2001), September, Nr. 5-6, 776–779. http://dx.doi.org/10.1007/s002530100708. – DOI 10.1007/s002530100708. – ISSN 0175–7598

[35] Haverkorn van Rijsewijk, Bart R B. ; Nanchen, Annik ; Nallet, Sophie ; Kleijn, Roelco J. ; Sauer, Uwe: Large-scale 13C-flux analysis reveals distinct transcriptional control of respiratory and fermentative metabolism in Escherichia coli. In: *Molecular systems biology* 7 (2011), März, 477. http://dx.doi.org/10.1038/msb.2011.9. – DOI 10.1038/msb.2011.9. – ISSN 1744–4292

[36] Heinrich, Reinhart ; Schuster, Stefan: *The Regulation Of Cellular Systems*. Springer, 1996. – 416 S. – ISBN 978–0412032615

[37] Herbert, D. ; Elsworth, R. ; Telling, R. C.: The Continuous Culture of Bacteria; a Theoretical and Experimental Study. (1956), S. 601–622

[38] Hildyard, John C W. ; Halestrap, Andrew P.: Identification of the mitochondrial pyruvate carrier in Saccharomyces cerevisiae. In: *The Biochemical journal* 374 (2003), Oktober, Nr. Pt 3, 607–11. http://dx.doi.org/10.1042/BJ20030995. – DOI 10.1042/BJ20030995. – ISSN 1470–8728

[39] Hinkle, Peter C.: P/O ratios of mitochondrial oxidative phosphorylation. In: *Biochimica et biophysica acta* 1706 (2005), Januar, Nr. 1-2, 1–11. http://dx.doi.org/10.1016/j.bbabio.2004.09.004. – DOI 10.1016/j.bbabio.2004.09.004. – ISSN 0006–3002

[40] Hohmann, Stefan ; Neves, Maria J. ; Valckx, Dirk ; Thevelein, Johan: Evidence for trehalose-6-phosphate-dependent and -independent mechanisms in the control of sugar influx into yeast glycolysis. In: *Molecular Microbiology* 20 (1996), S. 981 –991

[41] Holzhütter, H G. ; Schuster, R ; Buckwitz, D ; Jacobasch, G: Mathematical modelling of metabolic pathways affected by an enzyme deficiency. In: *Biomedica biochimica acta* 49 (1990), Januar, Nr. 8-9, 791–800. http://www.ncbi.nlm.nih.gov/pubmed/2082922. – ISSN 0232–766X

[42] Holzhütter, Hermann-Georg: The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks. In: *European journal of biochemistry / FEBS* 271 (2004), Juli, Nr. 14, 2905–22. http://dx.doi.org/10.1111/j.1432-1033.2004.04213.x. – DOI 10.1111/j.1432–1033.2004.04213.x. – ISSN 0014–2956

[43] Hynne, F ; Danø, S ; Sø rensen, P G.: Full-scale model of glycolysis in Saccharomyces cerevisiae. In: *Biophysical Chemistry* 94 (2001), Nr. 1-2, 121–163. http://www.ncbi.nlm.nih.gov/pubmed/11744196

[44] Ibarguren, Izaskun ; Diaz-Enrich, Maria J. ; Cao, Jesús ; Fernandez, Montserrat ; Barcia, Ramiro ; Villamarin, José A. ; Ramos-Martinez, Juán I.: Regulation of the futile cycle of fructose phosphate in sea mussel. In: *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* 126 (2000), August, Nr. 4, 495–501. http://dx.doi.org/10.1016/S0305-0491(00)00211-X. – DOI 10.1016/S0305–0491(00)00211–X. – ISSN 10964959

[45] Ideker, T ; Galitski, T ; Hood, L: A new approach to decoding life: systems biology. In: *Annual review of genomics and human genetics* 2 (2001), Januar, 343–72. http://dx.doi.org/10.1146/annurev.genom.2.1.343. – DOI 10.1146/annurev.genom.2.1.343. – ISSN 1527–8204

[46] Jong-Gubbels, P de ; Bauer, J ; Niederberger, P ; Stückrath, I ; Kötter, P ; Dijken, J P. ; Pronk, J T.: Physiological characterisation of a pyruvate-carboxylase-negative Saccharomyces cerevisiae mutant in batch and chemostat cultures. In: *Antonie van Leeuwenhoek* 74 (1998), November, Nr. 4, 253–63. http://www.ncbi.nlm.nih.gov/pubmed/10081585. – ISSN 0003–6072

[47] Jouhten, Paula ; Rintala, Eija ; Huuskonen, Anne ; Tamminen, Anu ; Toivari, Mervi ; Wiebe, Marilyn ; Ruohonen, Laura ; Penttilä, Merja ; Maaheimo, Hannu: Oxygen dependence of metabolic fluxes and energy generation of Saccharomyces cerevisiae CEN.PK113-1A. In: *BMC systems biology* 2 (2008), Januar, 60. http://dx.doi.org/10.1186/1752-0509-2-60. – DOI 10.1186/1752–0509–2–60. – ISSN 1752–0509

[48] Kanehisa, Minoru ; Goto, Susumu ; Furumichi, Miho ; Tanabe, Mao ; Hirakawa, Mika: KEGG for representation and analysis of molecular networks involving diseases and drugs. In: *Nucleic acids research* 38 (2010), Januar, Nr. Database issue, D355–60. http://dx.doi.org/10.1093/nar/gkp896. – DOI 10.1093/nar/gkp896. – ISSN 1362–4962

[49] Kanehisa, Minoru ; Goto, Susumu ; Hattori, Masahiro ; Aoki-Kinoshita, Kiyoko F. ; Itoh, Masumi ; Kawashima, Shuichi ; Katayama, Toshiaki ; Araki, Michihiro ; Hirakawa, Mika: From genomics to chemical genomics: new developments in KEGG. In: *Nucleic acids research* 34 (2006), Januar, Nr. Database issue, D354–7. http://dx.doi.org/10.1093/nar/gkj102. – DOI 10.1093/nar/gkj102. – ISSN 1362–4962

[50] Klipp, Edda ; Herwig, Ralf ; Kowald, Axel ; Wierling, Christoph ; Lehrach, Hans: *Systems Biology in Practice*. Wiley-VCH, 2005. – 465 S.

[51] Kovárová-Kovar, Karin ; Egli, Thomas: Growth Kinetics of Suspended Microbial Cells: From Single- Substrate-Controlled Growth to Mixed-Substrate Kinetics. In: *Society* 62 (1998), Nr. 3, S. 646–666

[52] Kretschmer, M ; Schellenberger, W ; Otto, A ; Kessler, R ; Hofmann, E: Fructose-2,6-bisphosphatase and 6-phosphofructo-2-kinase are separable in yeast. In: *The Biochemical journal* 246 (1987), September, Nr. 3, 755–9. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1148341&tool=pmcentrez&rendertype=abstract. – ISSN 0264–6021

[53] Kuile, B H. ; Westerhoff, H V.: Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. In: *FEBS letters* 500 (2001), Juli, Nr. 3, 169–71. http://www.ncbi.nlm.nih.gov/pubmed/11445079. – ISSN 0014–5793

[54] Lee, M V. ; Topper, Scott E. ; Hubler, Shane L. ; Hose, James ; Wenger, Craig D. ; Coon, Joshua J. ; Gasch, Audrey P.: A dynamic model of proteome changes reveals new roles for transcript alteration in yeast. In: *Molecular Systems Biology* 7 (2011), Juli. http://dx.doi.org/10.1038/msb.2011.48. – DOI 10.1038/msb.2011.48. – ISSN 1744–4292

[55] Lemuth, K ; Hardiman, T ; Winter, S ; Pfeiffer, D ; Keller, M A. ; Lange, S ; Reuss, M ; Schmid, R D. ; Siemann-Herzberg, M: Global transcription and metabolic flux analysis of Escherichia coli in glucose-limited fed-batch cultivations. In: *Applied and environmental microbiology* 74 (2008), November, Nr. 22, 7002–15. http://dx.doi.org/10.1128/AEM.01327-08. – DOI 10.1128/AEM.01327–08. – ISSN 1098–5336

[56] Maaheimo, H ; Fiaux, J ; Cakar, Z P. ; Bailey, J E. ; Sauer, U ; Szyperski, T: Central carbon metabolism of Saccharomyces cerevisiae explored by biosynthetic fractional (13)C labeling of common amino acids. In: *European journal of biochemistry / FEBS* 268 (2001), April, Nr. 8, 2464–79. http://www.ncbi.nlm.nih.gov/pubmed/11298766. – ISSN 0014–2956

[57] Mescam, Muriel ; Vinnakota, Kalyan C. ; Beard, Daniel a.: Identification of the catalytic mechanism and estimation of kinetic parameters for fumarase. In: *The Journal of biological chemistry* 286 (2011), Juni, Nr. 24, 21100–9. http://dx.doi.org/10.1074/jbc.M110.214452. – DOI 10.1074/jbc.M110.214452. – ISSN 1083–351X

[58] Monod, J: La technique de la culture continue, théorie et applications. In: *Annales de Institute Pasteur Paris* 79 (1950), Nr. 4, S. 390–410

[59] Nielsen, Jens ; Villadsen, John ; Lidén, Gunnar: *Bioreaction Engineering Principles*. 2. Springer, 2002. – 540 S. – ISBN 0306473496

[60] Notebaart, Richard a. ; Teusink, Bas ; Siezen, Roland J. ; Papp, Balázs: Co-regulation of metabolic genes is better explained by flux coupling than by network distance. In: *PLoS computational biology* 4 (2008), Januar, Nr. 1, e26. http://dx.doi.org/10.1371/journal.pcbi.0040026. – DOI 10.1371/journal.pcbi.0040026. – ISSN 1553–7358

[61] Novick, A. ; Szilard, L.: Experiments with the chemostat on spontaneous mutations of bacteria. In: *Proceedings of the National Academy of Sciences of the United States of America* 36 (1950), Nr. 12, 708. http://dx.doi.org/PMC1063276. – DOI PMC1063276

[62] Novick, Aaron ; Szilard, Leo: Description of the Chemostat. In: *Science* 112 (1950), S. 715–716

[63] Piper, Matthew D W. ; Daran-Lapujade, Pascale ; Bro, Christoffer ; Regenberg, Birgitte ; Knudsen, Steen ; Nielsen, Jens ; Pronk, Jack T.: Reproducibility of oligonucleotide microarray transcriptome analyses. An interlaboratory comparison using chemostat cultures of Saccharomyces cerevisiae. In: *The Journal of biological chemistry* 277 (2002), Oktober, Nr. 40, 37001–8. http://dx.doi.org/10.1074/jbc.M204490200. – DOI 10.1074/jbc.M204490200. – ISSN 0021–9258

[64] Pitkänen, JP: *Impact of Xylose and Mannose on Central Metabolism of Yeast Saccharomyces cerevisiae (PhD thesis)*, Diss., 2005. http://lib.tkk.fi/Diss/2005/isbn9512278944/

[65] Postma, E ; Kuiper, A ; Tomasouw, W F. ; Scheffers, W A. ; Dijken, J P.: Competition for Glucose between the Yeasts Saccharomyces cerevisiae and Candida utilis. In: *Applied and environmental microbiology* 55 (1989), Nr. 12, S. 3214–3220

[66] Ralser, Markus ; Wamelink, Mirjam M. ; Kowald, Axel ; Gerisch, Birgit ; Heeren, Gino ; Struys, Eduard A. ; Klipp, Edda ; Jakobs, Cornelis ; Breitenbach, Michael ; Lehrach, Hans ; Krobitsch, Sylvia: Dynamic rerouting of the carbohydrate flux is key to counteracting oxidative stress. In: *Journal of biology* 6 (2007), Januar, Nr. 4, 10. http://dx.doi.org/10.1186/jbiol61. – DOI 10.1186/jbiol61. – ISSN 1475–4924

[67] Rapoport, T ; Heinrich, Reinhart: Mathematical analysis of multienzyme systems. I. Modelling of the glycolysis of human erythrocytes. In: *Biosystems* 7 (1975), Juli, Nr. 1, 120–129. http://dx.doi.org/10.1016/0303-2647(75)90049-0. – DOI 10.1016/0303–2647(75)90049–0. – ISSN 03032647

[68] Rapoport, T a. ; Heinrich, Reinhart ; Rapoport, S M.: The regulatory principles of glycolysis in erythrocytes in vivo and in vitro. A minimal comprehensive model describing steady states, quasi-steady states and time-dependent processes. In: *The Biochemical journal* 154 (1976), Februar, Nr. 2, 449–69. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1172726&tool=pmcentrez&rendertype=abstract. – ISSN 0264–6021

[69] Rintala, Eija ; Toivari, Mervi ; Pitkänen, Juha-Pekka ; Wiebe, Marilyn G. ; Ruohonen, Laura ; Penttilä, Merja: Low oxygen levels as a trigger for enhancement of respiratory metabolism in Saccharomyces cerevisiae. In: *BMC genomics* 10 (2009), Januar, 461. http://dx.doi.org/10.1186/1471-2164-10-461. – DOI 10.1186/1471–2164–10–461. – ISSN 1471–2164

[70] Rintala, Eija ; Wiebe, Marilyn G. ; Tamminen, Anu ; Ruohonen, Laura ; Penttilä, Merja: Transcription of hexose transporters of Saccharomyces cerevisiae is affected by change in oxygen provision. In: *BMC microbiology* 8 (2008), Januar, 53. http://dx.doi.org/10.1186/1471-2180-8-53. – DOI 10.1186/1471–2180–8–53. – ISSN 1471–2180

[71] Rizzi, M ; Baltes, M ; Theobald, U ; Reuss, M: In vivo analysis of metabolic dynamics in Saccharomyces cerevisiae: II. Mathematical model. In: *Biotechnology and bioengineering* 55 (1997), August, Nr. 4, 592–608. http://dx.doi.org/10.1002/(SICI)1097-0290(19970820)55:4<592::AID-BIT2>3.0.CO;2-C. – DOI 10.1002/(SICI)1097–0290(19970820)55:4<592::AID–BIT2>3.0.CO;2–C. – ISSN 0006–3592

[72] Rizzi, M ; Theobald, U ; Querfurth, E ; Rohrhirsch, T ; Baltes, M ; Reuss, M: In vivo investigations of glucose transport in Saccharomyces cerevisiae. In: *Biotechnology and bioengineering* 49 (1996), Februar, Nr. 3, 316–27. http://dx.doi.org/10.1002/(SICI)1097-0290(19960205)49:3<316::AID-BIT10>3.0.CO;2-C. – DOI 10.1002/(SICI)1097–0290(19960205)49:3<316::AID–BIT10>3.0.CO;2–C. – ISSN 0006–3592

[73] Rossell, Sergio ; Weijden, Coen C. d. ; Lindenbergh, Alexander ; Tuijl, Arjen van ; Francke, Christof ; Bakker, Barbara M. ; Westerhoff, Hans V.: Unraveling the complexity of flux regulation: a new method demonstrated for nutrient starvation in Saccharomyces cerevisiae. In: *Proceedings of the National Academy of Sciences of the United States of America* 103 (2006), Februar, Nr. 7, 2166–71. http://dx.doi.org/10.1073/pnas.0509831103. – DOI 10.1073/pnas.0509831103. – ISSN 0027–8424

[74] Sauer, Uwe: *pers. comm., e-mail (10. oct. 2011)*

[75] Sauer, Uwe: *pers. comm., email (24 june 2011)*

[76] Scheer, Maurice ; Grote, Andreas ; Chang, Antje ; Schomburg, Ida ; Munaretto, Cornelia ; Rother, Michael ; Söhngen, Carola ; Stelzer, Michael ; Thiele, Juliane ; Schomburg, Dietmar: BRENDA, the enzyme information system in 2011. In: *Nucleic acids research* 39 (2011), Januar, Nr. Database issue, D670–6. http://dx.doi.org/10.1093/nar/gkq1089. – DOI 10.1093/nar/gkq1089. – ISSN 1362–4962

[77] Schmidt, H ; Jirstrand, M: Systems Biology Toolbox for MATLAB: a computational platform for research in systems biology. In: *Bioinformatics* 22 (2006), Nr. 4, 514–515. http://dx.doi.org/10.1093/bioinformatics/bti799. – DOI 10.1093/bioinformatics/bti799. – ISSN 13674803

[78] Schuetz, Robert ; Kuepfer, Lars ; Sauer, Uwe: Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli. In: *Molecular systems biology* 3 (2007), Januar, Nr. 119, 119. http://dx.doi.org/10.1038/msb4100162. – DOI 10.1038/msb4100162. – ISSN 1744–4292

[79] Sel'kov, E E. ; Avseenko, N V. ; Kirsta, Iu B.: [Allosteric regulation in the open futile cycle fructose-6-P–fructose-1,6-P2]. In: *Biofizika* 24, Nr. 5, 829–35. http://www.ncbi.nlm.nih.gov/pubmed/39635. – ISSN 0006–3029

[80] Silverman, Sanford J. ; Petti, Allegra a. ; Slavov, Nikolai ; Parsons, Lance ; Briehof, Ryan ; Thiberge, Stephan Y. ; Zenklusen, Daniel ; Gandhi, Saumil J. ; Larson, Daniel R. ; Singer, Robert H. ; Botstein, David: Metabolic cycling in single yeast cells from unsynchronized steady-state populations limited on glucose or phosphate. In: *Proceedings of the National Academy of Sciences of the United States of America* 107 (2010), April, Nr. 15, 6946–51. http://dx.doi.org/10.1073/pnas.1002422107. – DOI 10.1073/pnas.1002422107. – ISSN 1091–6490

[81] Smith, John M. ; Szathmáry, Eörs: *The Major Transitions in Evolution*. Oxford : Oxford University Press, 1995. – 346 S. – ISBN 019850294X

[82] Sprague, George F. ; Winans, Stephen C.: Eukaryotes learn how to count: quorum sensing by yeast. In: *Genes & development* 20 (2006), Mai, Nr. 9, 1045–9. http://dx.doi.org/10.1101/gad.1432906. – DOI 10.1101/gad.1432906. – ISSN 0890–9369

[83] Stadtman, E R. ; Chock, P B.: Superiority of interconvertible enzyme cascades in metabolic regulation: Analysis of monocyclic systems*. 74 (1977), Nr. 7, S. 2761–2765

[84] Staples, James F. ; Koen, Erin L. ; Laverty, Terence M.: 'Futile cycle' enzymes in the flight muscles of North American bumblebees. In: *The Journal of experimental biology* 207 (2004), Februar, Nr. Pt 5, 749–54. http://www.ncbi.nlm.nih.gov/pubmed/14747407. – ISSN 0022–0949

[85] Stephanopoulos, Gregory N. ; Aristidou, Aristos A. ; Nielsen, Jens: *Metabolic Engineering: Principles and Methodologies*. Academic Press, 1998. – 725 S. – ISBN 0126662606

[86] Sugden, MaryÂ C. ; Holness, MarkÂ J.: Trials, tribulations and finally, a transporter: the identification of the mitochondrial pyruvate transporter. In: *Biochemical Journal* 374 (2003), September, Nr. 3, e1. http://dx.doi.org/10.1042/BJ20031105. – DOI 10.1042/BJ20031105. – ISSN 02646021

[87] Teusink, B ; Passarge, J ; Reijenga, C A. ; Esgalhado, E ; van Der Weijden, C C. ; Schepper, M ; Walsh, M C. ; Bakker, B M. ; Dam, K van ; Westerhoff, H V. ; Snoep, J L.: Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. In: *European journal of biochemistry / FEBS* 267 (2000), September, Nr. 17, 5313–29. http://www.ncbi.nlm.nih.gov/pubmed/10951190. – ISSN 0014–2956

[88] Teusink, B ; Walsh, M C. ; Dam, K van ; Westerhoff, H V.: The danger of metabolic pathways with turbo design. In: *Trends in biochemical sciences* 23 (1998), Mai, Nr. 5, 162–9. http://www.ncbi.nlm.nih.gov/pubmed/9612078. – ISSN 0968–0004

[89] Teusink, Bastiaan: *Exposing a complex metabolic system: glycolysis in Saccharomyces cerevisiae*, University of Amsterdam, Diss., 1999. – 231 S.

[90] Theobald, U ; Mailinger, W ; Baltes, M ; Rizzi, M ; Reuss, M: In vivo analysis of metabolic dynamics in Saccharomyces cerevisiae : I. Experimental observations. In: *Biotechnology and bioengineering* 55 (1997), Juli, Nr. 2, 305–16. http://dx.doi.org/10.1002/(SICI)1097-0290(19970720)55:2<305::AID-BIT8>3.0.CO;2-M. – DOI 10.1002/(SICI)1097–0290(19970720)55:2<305::AID–BIT8>3.0.CO;2–M. – ISSN 0006–3592

[91] Theobald, U ; Mailinger, W ; Reuss, M ; Rizzi, M: In vivo analysis of glucose-induced fast changes in yeast adenine nucleotide pool applying a rapid sampling technique. In: *Analytical biochemistry* 214 (1993), Oktober, Nr. 1, 31–7. http://dx.doi.org/10.1006/abio.1993.1452. – DOI 10.1006/abio.1993.1452. – ISSN 0003–2697

[92] Thevelein, Johan M. ; Hohmann, Stefan: Trehalose synthase: guard to the gate of glycolysis in yeast? In: *Trends in Biochemical Sciences* 20 (1995), Januar, Nr. 1, 3–10. http://dx.doi.org/10.1016/S0968-0004(00)88938-0. – DOI 10.1016/S0968–0004(00)88938–0. – ISSN 09680004

[93] Vaseghi, S ; Baumeister, a ; Rizzi, M ; Reuss, M: In vivo dynamics of the pentose phosphate pathway in Saccharomyces cerevisiae. In: *Metabolic engineering* 1 (1999), April, Nr. 2, 128–40. http://dx.doi.org/10.1006/mben.1998.0110. – DOI 10.1006/mben.1998.0110. – ISSN 1096–7176

[94] Verduyn, C ; Postma, E ; Scheffers, W A. ; Van Dijken, J P.: Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. In: *Yeast (Chichester, England)* 8 (1992), Juli, Nr. 7, 501–17. http://dx.doi.org/10.1002/yea.320080703. – DOI 10.1002/yea.320080703. – ISSN 0749–503X

[95] Voet, D ; Voet, J G.: *Biochemistry, 3rd Edition*. Wiley, 2004 http://pubs.acs.org/doi/abs/10.1021/bi00423a001

[96] Weusthuis, R A. ; Pronk, J T. ; Broek, P J. d. ; Dijken, J P.: Chemostat cultivation as a tool for studies on sugar transport in yeasts. In: *Microbiological reviews* 58 (1994), Dezember, Nr. 4, 616–30. http://dx.doi.org/10.1007/BF00016496. – DOI 10.1007/BF00016496. – ISSN 0146–0749

[97] Wiebe, Marilyn G. ; Rintala, Eija ; Tamminen, Anu ; Simolin, Helena ; Salusjärvi, Laura ; Toivari, Mervi ; Kokkonen, Juha T. ; Kiuru, Jari ; Ketola, Raimo a. ; Jouhten, Paula ; Huuskonen, Anne ; Maaheimo, Hannu ; Ruohonen, Laura ; Penttilä, Merja: Central carbon metabolism of Saccharomyces cerevisiae in anaerobic, oxygen-limited and fully aerobic steady-state conditions and following a shift to anaerobic conditions. In: *FEMS yeast research* 8 (2008), Februar, Nr. 1, 140–54. http://dx.doi.org/10.1111/j.1567-1364.2007.00234.x. – DOI 10.1111/j.1567–1364.2007.00234.x. – ISSN 1567–1356

[98] Wolber, Paul K. ; Collins, Patrick J. ; Lucas, Anne B. ; De Witte, Anniek ; Shannon, Karen W.: The Agilent in situ-synthesized microarray platform. In: *Methods in enzymology* 410 (2006), Januar, 28–57. http://dx.doi.org/10.1016/S0076-6879(06)10002-6. – DOI 10.1016/S0076–6879(06)10002–6. – ISSN 0076–6879

[99] Wong, Matthew S. ; Raab, R M. ; Rigoutsos, Isidore ; Stephanopoulos, Gregory N. ; Kelleher, Joanne K.: Metabolic and transcriptional patterns accompanying glutamine depletion and repletion in mouse hepatoma cells: a model for physiological regulatory networks. In: *Physiological genomics* 16 (2004), Januar, Nr. 2, 247–55. http://dx.doi.org/10.1152/physiolgenomics.00088.2003. – DOI 10.1152/physiolgenomics.00088.2003. – ISSN 1531–2267

[100] Zaim, Jolanta ; Speina, Elzbieta ; Kierzek, Andrzej M.: Identification of new genes regulated by the Crt1 transcription factor, an effector of the DNA damage checkpoint pathway in Saccharomyces cerevisiae. In: *The Journal of biological chemistry* 280 (2005), Januar, Nr. 1, 28–37. http://dx.doi.org/10.1074/jbc.M404669200. – DOI 10.1074/jbc.M404669200. – ISSN 0021–9258

[101] Zamamiri, A Q. ; Birol, G ; Hjortsø, M A.: Multiple stable states and hysteresis in continuous, oscillating cultures of budding yeast. In: *Biotechnology and bioengineering* 75 (2001), November, Nr. 3, 305–12. http://www.ncbi.nlm.nih.gov/pubmed/11590603. – ISSN 0006–3592

Selbständigkeitserklärung

Ich erkläre hiermit, diese vorliegende Arbeit selbstsändig und nur unter Verwendung der angegebenen Hilfsmittel und Literatur verfaßt zu haben.

József Bruck